ICONS 2022

The Seventeenth International Conference on Systems

ISBN: 978-1-61208-941-6

April 24 - 28, 2022

Barcelona, Spain

ICONS 2022 Editors

Mo Mansouri, Stevens Institute of Technology, USA, University of South-Eastern Norway, Norway, USA

ICONS 2022

Forward

The Seventeenth International Conference on Systems (ICONS 2022) continued a series of events covering a broad spectrum of topics including fundamentals on designing, implementing, testing, validating and maintaining various kinds of software and hardware systems.

In the last years, new system concepts have been promoted and partially embedded in new deployments. Anticipative systems, autonomic and autonomous systems, self-adapting systems, or on-demand systems are systems exposing advanced features. These features demand special requirements specification mechanisms, advanced behavioral design patterns, special interaction protocols, and flexible implementation platforms. Additionally, they require new monitoring and management paradigms, as self-protection, self-diagnosing, self-maintenance become core design features.

The design of application-oriented systems is driven by application-specific requirements that have a very large spectrum. Despite the adoption of uniform frameworks and system design methodologies supported by appropriate models and system specification languages, the deployment of application-oriented systems raises critical problems. Specific requirements in terms of scalability, real-time, security, performance, accuracy, distribution, and user interaction drive the design decisions and implementations. This leads to the need for gathering application-specific knowledge and develop particular design and implementation skills that can be reused in developing similar systems.

Validation and verification of safety requirements for complex systems containing hardware, software and human subsystems must be considered from early design phases. There is a need for rigorous analysis on the role of people and process causing hazards within safety-related systems; however, these claims are often made without a rigorous analysis of the human factors involved. Accurate identification and implementation of safety requirements for all elements of a system, including people and procedures become crucial in complex and critical systems, especially in safety-related projects from the civil aviation, defense health, and transport sectors.

Fundamentals on safety-related systems concern both positive (desired properties) and negative (undesired properties) aspects. Safety requirements are expressed at the individual equipment level and at the operational-environment level. However, ambiguity in safety requirements may lead to reliable unsafe systems. Additionally, the distribution of safety requirements between people and machines makes difficult automated proofs of system safety. This is somehow obscured by the difficulty of applying formal techniques (usually used for equipment related safety requirements) to derivation and satisfaction of human-related safety requirements (usually, human factor techniques are used).

We take here the opportunity to warmly thank all the members of the ICONS 2022 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to ICONS 2022. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the ICONS 2022 organizing committee for their help in handling the logistics of this event.

ICONS 2022 Chairs

ICONS 2022 Steering Committee

David Inkermann, Technische Universität Clausthal, Institute of Mechanical Engineering, Germany

ICONS 2022
Committee

ICONS 2022 Steering Committee

David Inkermann, Technische Universität Clausthal, Institute of Mechanical Engineering, Germany
Christoph Knieke, Technische Universität Clausthal, Institute for Software and Systems Engineering, Germany
Giulio Telleschi, MBDA, Italy
Mo Mansouri, Stevens Institute of Technology, USA, University of South-Eastern Norway, Norway
Mark Austin, University of Maryland, USA
Gerrit Muller, University of South-Eastern Norway, Norway, TNO - Eindhoven, The Netherlands

ICONS 2022 Publicity Chairs

Javier Rocher, Universitat Politècnica de València, Spain
Lorena Parra, Universitat Politècnica de València, Spain

ICONS 2022 Technical Program Committee

Qammer H. Abbasi, University of Glasgow, Scotland, UK
Mohamed A. Abd El Ghany, German University in Cairo, Egypt
Witold Abramowicz, Poznan University of Economics, Poland
Afaq Ahmad, Sultan Qaboos University, Oman
Abdelouhab Aitouche, YNCREA/HEI | University of Lille, France
Ali Al-Humairi, German University of Technology (GUtech), Oman
Walid Al-Hussaibi, Southern Technical University (STU), Iraq
Mohammed Al-Khafajiy, University of Reading, UK
Saloua Bel Hadj Ali, University of Tunis-El Manar / University of Gabès, Tunisia
Marc Austin, University Of Maryland, USA
Muhammed Ali Aydin, Istanbul University-Cerrahpasa, Turkey
Snježana Babić, Juraj Dobrila University of Pula, Croatia
Lubomir Bakule, Institute of Information Theory and Automation, Czech Republic
Kamel Barkaoui, Cedric - Cnam, Paris, France
Janibul Bashir, National Institute of Technology, Srinagar, India
Suvadip Batabyal, BITS Pilani, Hyderabad Campus, India
Alejandro J. Bianchi, LIVEWARE S.A. / Universidad Catolica Argentina, Argentina
Francesco Bianconi, Università degli Studi di Perugia, Italy
Birthe Boehm, Siemens AG, Germany
Sander Bohte, Machine Learning group - CWI, Amsterdam, The Netherlands
Rahma Boucetta, University of Sfax, Tunisia
Frédéric Bousefsaf, Université de Lorraine, France
Antonio Brogi, University of Pisa, Italy
Eugenio Brusa, Politecnico di Torino, Italy
Gert Cauwenberghs, University of California, San Diego, USA
Rachid Chelouah, ETIS UMRS CNRS laboratory | CY Cergy Paris University, France
Dejiu Chen, KTH, Sweden
Albert M. K. Cheng, University of Houston, USA
Larbi Chrifi-Alaoui, University of Picardie Jules Verne, France
François Coallier, École de technologie supérieure, Montreal, Canada

Copyright Information

Table of Contents

The Systems Approach to Rethinking the Gender-Equality Paradox in STEM in the Context of Norwegian Educational System and Workforce

Evelina Anita Iversen
Department of Science and Industry systems
University of South-Eastern Norway
Kongsberg, Norway
evelina.a.iversen@gmail.com

Mo Mansouri
Department of Science and Industry systems
University of South-Eastern Norway
Kongsberg, Norway
Mo.Mansouri@usn.no

***Abstract*—Gender equality in Science, Technology, Engineering, and Mathematics (STEM) has been an ongoing concern of policymakers and researchers. The equality issues of socioeconomic systems are inherently complex due to their dynamic nature and many tightly coupled variables. These initial conditions of the systemic gender issue inspired the authors of this paper to apply the systems multistakeholder approach to this topic. The goal is to show that it is possible to assess and organise the dynamic behaviour of the agents influencing the mechanics of women's choices. The authors chose Casual Loop Diagrams (CLD) as a visualisation tool. Firstly, the paper focuses on socioeconomic factors that affect adolescent girls' willingness and ability to pursue a STEM-related education and career. It has been shown that the problem is multifaceted; hence, the decision has been made to focus on the mechanisms that influence college enrollment. Furthermore, the authors considered the geopolitical context with Norway in focus. The paper also outlines the choice dynamic and suggests that relatable female representation can increase gender parity in STEM. Finally, it has been shown that CLD might be a useful modelling tool for the multidisciplinary team working on solving systemic issues. CLD has the capability of bringing a middle ground between policymakers, neurologists, social scientists and other specialists required to analyse the issue thoroughly.**

***Index Terms*—Complex systems, education, gender studies, STEM.**

I. INTRODUCTION

One of the most prominent intentions of The Sustainable Development Goals is to promote fighting inequality and discrimination all around the globe. This issue is specifically targeted in *"Goal 5 - gender equality"*. There, among other ambitions, is to *"ensure women's full participation [...] in decision-making"* [4]. Since engineers and other representatives of STEM-related workforce play a great role in the shaping of humanity's future by making decisions on technology in *"a response to societal needs"* [2], it is worth analysing women's representation in this social group. One could argue that isolating a block of society to perform gender-related analysis can narrow the scope of issues of inequality.

Nevertheless, since human interaction has a high level of complexity, it is rather unlikely that isolation would eliminate the effects of society as a whole. Therefore, to reduce the complexity of the system and the issue discussed in this paper, it can be a good starting point to focus on one social group. First, however, close attention must be paid to the problem's cultural, socioeconomic, and geographic factors. That will be the main focus of Section I. The rest of this paper is organised as follows. Section II illustrates the modelling process of the mechanics of choice associated with STEM-related higher education. In Section III, the authors elaborate on the relationships visualised in the previous section while indicating a possible intervention. In conclusion, Section IV addresses the value of the systems multistakeholder approach and covers suggestions for further work.

A. Women's under-representation

Since resolving women's under-representation in STEM-related education and workforce is an issue that is often brought up on a systemic level, one could argue that it is evident that governments and industries are the stakeholders of this problem. Indeed, it is evident that a larger pool of trained professionals will positively affect the economy. That is because engineering and science, in general, are rapidly developing fields, which are often understaffed [3].

Educational institutions and research centres are motivated to employ more women since they bring their unique perspective to the projects. As an example, one could mention a rather famous *"Seat belt design case"*, where male researchers created the guidelines, which did not take into consideration gender-influenced anatomical differences [1].

On the other hand, some groups would oppose introducing more women in STEM. Those groups might be promoting "conservative gender roles," where women either engage in unpaid labour in their households or are employed in traditionally female jobs. Other groups could be against female representation since that would increase competition, making it harder to receive higher achievements. It is difficult to dismiss the existence of these undeniably misogynistic opinions, even though they are less prominent in the geographical context of this paper compared to other countries (see I-B1).

 ISBN: 978-1-61208-941-6

1) Statistics in Norway: Statistisk Sentralbyrå (Statistics Norway) has recently published data on higher education enrollment in the country. The document [7] shows that 60.2% of all applicants in 2021 are women. However, they represent only 30.9% of all junior students at STEM faculties. It has also been mentioned that engineering educational institutions experienced a slight decrease in women's representation (−0.9%), balanced by a significant increase in maritime technology (+7%).

The recent statistics can be perceived as discouraging, particularly if one would take into consideration an article published by *Norges Ingeniør- og Teknologorganisasjon* (Norwegian Engineering and Technological Organisation) [6], where they call for more female representation and point out that only 26% of STEM-workplaces are taken by women.

B. Willingness and ability

Since pursuing a certain educational- and career path is guided by a choice of an individual, one can take freedom to condense factors contributing to that particular decision into two groups: individual's willingness and their ability (see Fig. 1).

One can assume that under-representation of women in

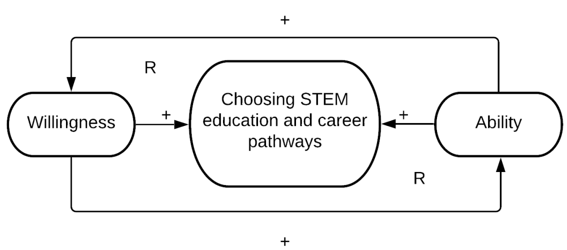


Fig. 1. Casual loop - best case scenario

STEM fields is due to their intellectual inability, rooted in biological inferiority. Even though one can find this point of view as admittedly false, Harvard University president Lawrence Summers once *"suggested that innate biological differences may help explain why men have more career success in science and mathematics than women"* [8]. This claim currently has little neurological support. Nevertheless, the statement can indicate the cultural atmosphere in some educational institutions. That atmosphere, which feeds on degrading stereotypes about women, is one of the reasons women are not willing to pursue a career in STEM, which reinforces male representation in the field (see Fig. 2). Fortunately, women's science performance concerning their male colleagues has been covered in other research articles that yielded contrary findings [9]. However, those studies alone cannot affect gender bias significantly enough to attract more women.

1) National gender-equality paradox: In 2018, Stoet and Geary [9] published an article where they analysed relative academic strengths of adolescent girls and boys concerning

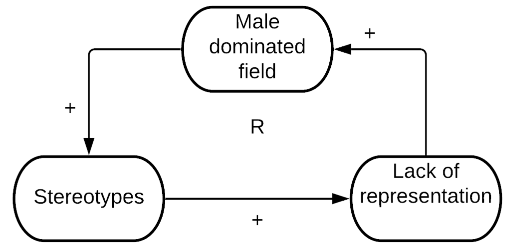


Fig. 2. Casual Loop - stereotype reinforcement

the probability of pursuing STEM educational path. It has

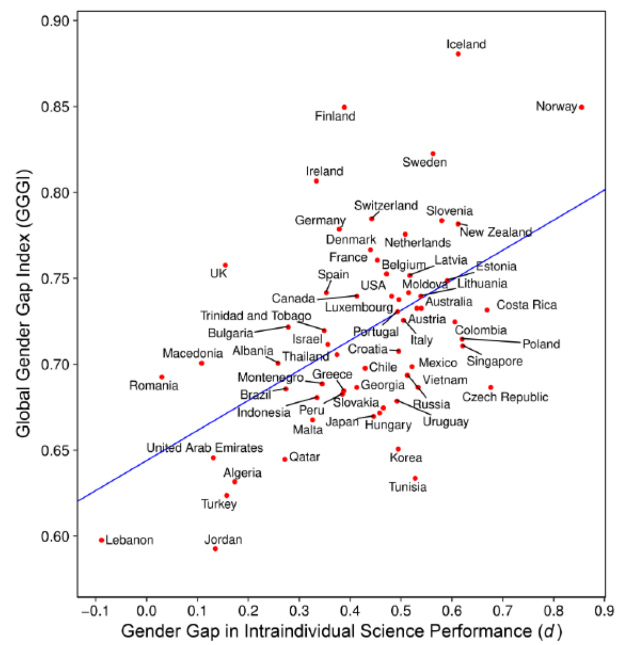


Fig. 3. Gender equality and sex differences vs intraindividual science performance

been found that in ≈ 66% of countries studied in the paper, girls performed at least similarly to boys. Additionally, *"in [...] all countries more girls [...] were capable of college-level STEM study than had enrolled"* [9]. To quantify gender parity, one often refers to Global Gender Gap Index (GGGI). In the countries where GGGI is high political, economic, and educational opportunities between men and women are relatively equal. In those countries, girls often outperform boys in sciences interindividual, but simultaneously are less ready to pursue STEM education-path (see Fig. 3 and Fig. 4).

Therefore a natural question arises: "if GGGI is high (their will is free) and women's intellectual abilities are similar to men's, why are women still under-represented in STEM in Norway?" One can argue that the analysis of this issue is three-sided. Firstly, one must understand why they are unwilling to pursue a STEM educational path; secondly, they are not fulfilling their studies and not following through with their

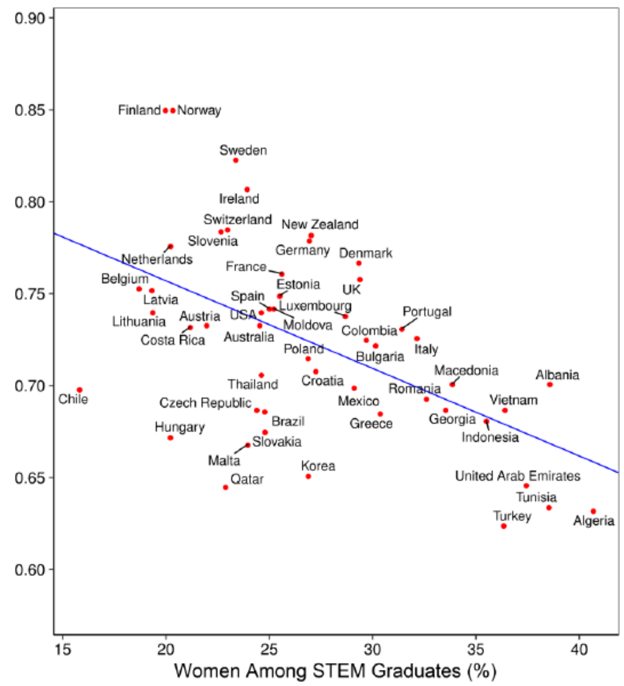


Fig. 4. Gender equality and sex differences vs women's prosperity relative to men

careers. Furthermore, lastly, why they change career paths. Since those three topics require isolated studies, the paper will continue with a case on college enrollment to provide a higher-quality analysis.

One then may attempt to structure information through Systems Thinking techniques and visualise the cause-result chains in a Casual Loop Diagram (CLD). A possible outcome of this procedure is represented in Section II together with reasoning behind factors chosen for dynamic assessment. Then, one can allocate loops that require intervention and propose a solution that benefits gender parity in STEM. The discussion and possible interventions can be found in Section III.

II. MODELING

Modelling the gender-equality in STEM in Norway was performed by relying on, among other resources, research published by Ida Marie Andersen on factors that influence the education pathway [5].

The analysis will be built on the further development of CLD represented in Fig. 1.

A. Parental influence

Ida Marie, in her paper, mentioned that parental influence is an essential factor of career and educational path choice for adolescents. The probability of a girl pursuing a STEM career is higher if she has a caregiver that has a job in that field. However, it is essential to point out that this analysis does not consider the relationship issues between a girl and her caregiver. Hence, in the best-case scenario, a caregiver or a parent increases the willingness of their child to pursue a career in STEM through encouraging and personal example. Then, engagement of a child will excite a caregiver, creating a reinforcement loop (Fig. 5).

Parental influence also affects a girl's academic performance, since they probably can help with homework and contribute to the general understanding of STEM-related syllabus, thus promoting the ability of a girl to choose a technical path in the future (see Fig. 6).

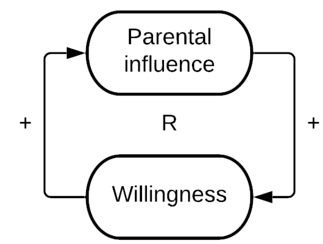


Fig. 5. Parental influence (willingness) - reinforcement loop

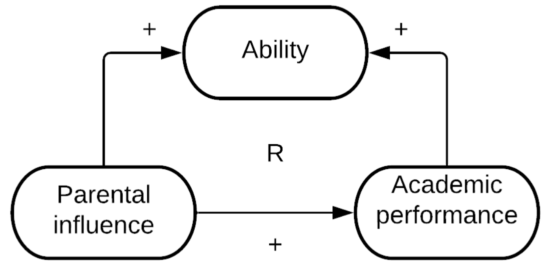


Fig. 6. Parental influence (ability) - reinforcement loop

1) Immigrant background: The work in [5] shows that parents who emigrated to Norway from other (non-Scandinavian) countries have an additional influence on their children: boys and girls. That point is supported by [9], where girls from countries with lesser general gender equality than in Norway are represented in STEM programs almost as often as boys.

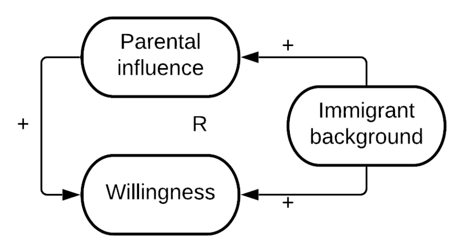


Fig. 7. Immigrant background (parents) - reinforcement loop

 ISBN: 978-1-61208-941-6

B. Gender bias - the reinforcement loop

Previous topics and loops that have been covered showed mechanisms that can positively affect a girl's decision when choosing a STEM-related career. The issues that can affect that decision negatively are generally rooted in gender bias (see Fig. 8). Interestingly enough, this destructive tendency is found both in women and men in the field. It has also been found that peer pressure in the context of general upper secondary education in Norway can swing girls away from engineering and mathematics since a STEM career path is not surrounded with the same amount of prestige as, for example, studies in Business and Administration, or Healthcare (see Fig. 9). Unfortunately, girls' unwillingness to pursue STEM careers feeds into stereotypes, reinforcing gender bias and lack of representation, making it a vicious circle.

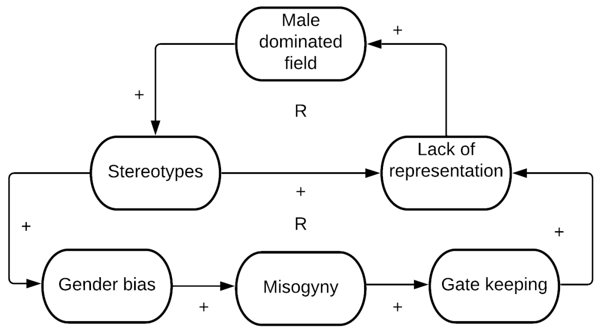


Fig. 8. Misogyny reinforcement loops

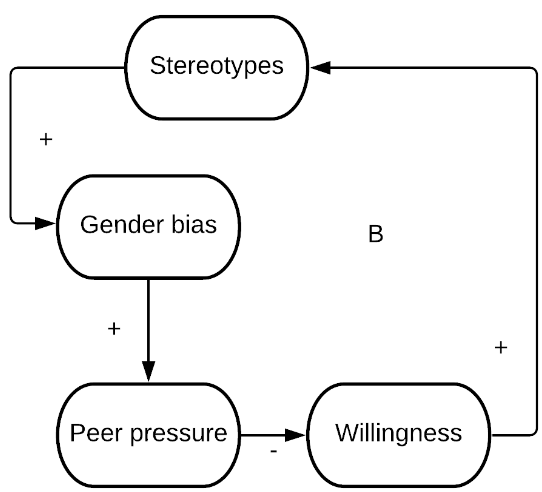


Fig. 9. Peer pressure prestige loop - balancing

C. Governmental support

The Norwegian government encourages women to enrol in STEM-related colleges and faculties by granting them additional points, thus, increasing their ability to compete for school places. It is important to point out that this support system is a part of the general gender-equality policy. This policy also gives men the same points for applying to female-dominant education paths. However, after interviewing some of their female colleagues in STEM, the author of this paper discovered that most of the women do not want to receive those points in the first place since that could suggest that they are less deserving of their current position than their male colleagues (see Fig. 10).

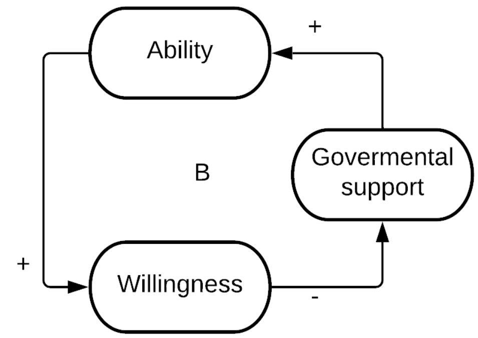


Fig. 10. Governmental policies - balance loop

III. DISCUSSION AND POSSIBLE INTERVENTIONS

It is possible to argue that fewer parameters negatively affect women's abilities than their willingness. Paradoxically, stereotypes and gender bias are at the root of those issues, even though Norway is considered one of the most gender-equal countries. That can be caused by a lack of comfort for women in the industry since there are very few success stories with female role models. And even if there is a success story, it is usually far beyond the reach of a "normal girl," since we more than often refer to Maria Skłodowska Curie or Valentina Tereshkova as an example of female triumph, which can be sometimes hard to associate with.

It is commonly known that usually, any hate, including misogyny, is fired by arrogance (willing and unwilling). One way to fight it is educating society on the topic by casting light on women studying and working in STEM *with comfort - without sacrificing their identity to a male-dominated field*. Additionally, one would want these women to be more relatable and reachable to adolescent girls than super-successful women in STEM, like Curie.

The authors propose that the positive dynamic can be encouraged by celebrating differences at workplaces and schools. Even though the industry is still overwhelmingly male-dominated, it can still be achieved even through minor changes. For example, one could end a practice of calling women in STEM "one of the boys" or expecting them to behave more masculinely not to be seen as an outcast. One could also encourage behaviours and activities that are usually

 ISBN: 978-1-61208-941-6

perceived as feminine. That would arguably reduce the effect of bias since male colleagues would have a chance to interact with women in their comfort zone and learn from them.

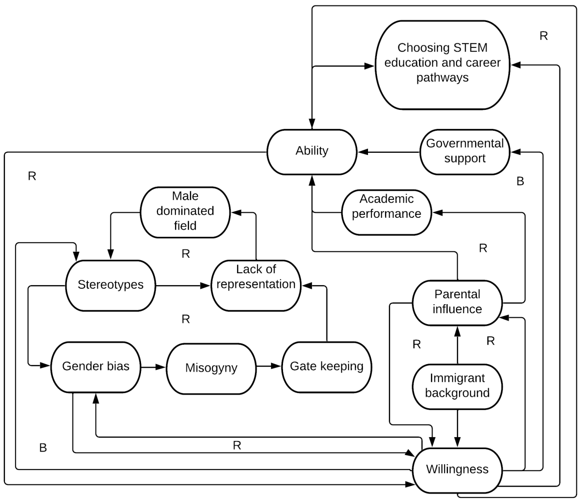


Fig. 11. Final Casual Loop Diagram

IV. CONCLUSION AND FUTURE WORK

Figure 11 shows the resulting CLD diagram, where one can see the attempt to visualise the mechanics of choice related to higher education in Norway. The resulting model demonstrates that complex socioeconomic issues can be structured understandably for specialists from different fields. That is a critical finding since continuing to build on the CLD STEM paradox model requires intervention from a multidisciplinary team with their knowledge of other agents affecting women's choice of educational path. In the future, one might reiterate existing CLD to learn more about the mechanisms of bias and gender stereotypes. That can include widening the model's scope and including some of the system's agents in the modelling process. Authors indicate that gradual holistic changes can be beneficial to make a STEM career path more attractive for women. For example, these interventions can focus on increasing the awareness of relatable female representation in the field.

REFERENCES

[1] A. Bert, ”3 reasons gender diversity is crucial to science”, Elsevier Connect, 2018. [Online]. Available: https://www.elsevier.com/connect/3-reasons-gender-diversity-is-crucial-to-science. [Accessed: 07-Mar-2022].

[2] H. Sladovich, Engineering as a social enterprise. Washington, DC: National Acad. Pr, 1993.

[3] M. Dahl, ”Continued high shortage of labor”, Metal Supply [Fortsatt høy mangel på arbeidskraft], 2022 . [Online]. Available: https://www.metalsupply.no/article /view/792477/fortsatt_hoy_mangel_pa_arbeidskraft. [Accessed: 07-Mar- 2022].

[4] Project Everyone, ”Goal 5: Gender equality - The Global Goals”, The Global Goals, 2022. [Online]. Available: https://www.globalgoals.org/goals/5-gender-equality/. [Accessed: 07- Mar- 2022].

[5] I. M. Andersen, ”Diversity in the Academy - Higher education in the face of a complex reality” [Mangfold i Akademia – Høyere utdanning i møte med en kompleks virkelighet], Oslo and Akershus University College, 2014. (HiOA Tema-utgivelser 2014, nr.2.)

[6] A. Berger, ”Women's Day March 8: - More girls should become engineers” [Kvinnedagen 8. mars: - Flere jenter bør bli ingeniør], NITO, 2022. [Online]. Available: https://www.nito.no/aktuelt/2022/3/vi-trenger-flere-kvinnelige-ingeniorer/. [Accessed: 07- Mar- 2022].

[7] Samordna Opptak, Hovedopptak til høyere utdanning ved universiteter og høgskoler juli 2021 [Main admission to higher education at universities and colleges July 2021], 2021.

[8] Nature Neuroscience, ”Separating science from stereotype”, Nat. Neurosci., vol. 8, no. 3, pp. 253-253, 2005.

[9] G. Stoet and D. Geary, ”The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education”, Psychological Science, vol. 29, no. 4, pp. 581-593, 2018.

A Systems Approach to Parking Assist System: Investigating Test and Verification Methodology

Celine Haugan Lia
University of South-Eastern Norway
USN
Kongsberg, Norway
Email: 238356@usn.no

Mo Mansouri
University of South-Eastern Norway
USN
Kongsberg, Norway
Email: Mo.Mansouri@usn.no

***Abstract*—This paper explores the parking assist system technology and how it has evolved over the years. It investigates the different types of subsystems and components that go into this small, yet complex system inside a vehicle and why this technology is needed. The system decomposition examines how everything is linked together and why all the subsystems and components are necessary for the system to function. The test and verification consider current test methods and explore existing equipment and software, especially for these cases. After discussing potential system requirements, a systems approach inclusive of both CATWOE (Customer, Actor, Transformation, World View, Owner, Environment) analysis and Systemigram methods is used to map out and get an overview of the system and its stakeholders. This paper presents a new approach to the way of looking at the parking assist system because taking all stakeholders and looking at the bigger picture is not considered in common practice. In a common analytical approach, engineers tend to focus on a separate part of the system without considering its interactions with the environment. Hence, looking at this problem from a systemic perspective is novel and constructive for this industry in the future.**

Keywords—Systems Thinking, PAS (Parking Assist System), ABS (Anti-lock Braking System), Systemigram, CATWOE.

I. Introduction

The roads are becoming increasingly busier and have been for many years and parking a vehicle in certain situations can be stressful no matter how experienced the driver is. Looking back in history, the first cars on the market were a lot smaller than the current design and were nowhere near as many vehicles on the road as we see today. Particularly in cities, the parking spaces can be tight and difficult to get into, however people are dependent on parking spaces. More people, more vehicles and tighter spaces unfortunately cause an increase in accidents. This happens not only on the roads, but in other places like parking lots and parking garages where the view around the vehicle is limited and the randomness of human behavior is responsible for a large percentage of accidents.

Parking assist technology has been around for many years and the first mainstream car to feature rear parking sensors was the Toyota Prius, although Toyota already had introduced ultrasonic back sensors as far back as 1982. The next technological development within the parking systems was the surrounding-view parking monitors/cameras, which Nissan first developed and introduced in 2007. The very first rear park assist system (semi-autonomous parking) was introduced in 2003 and has since then evolved into autonomous parking that is one of the latest features within parking technology [1].

II. Background and System Description

Looking into the modern car parking technology, one of the newest features that have been around for a few years is the semi-autonomous and autonomous parking assist system. The parking technology has evolved a lot over the last 20 years from when the rear parking sensors first were introduced. Today, most of the large car manufacturers provide parking assist technology in their new cars, or at least as a feature option that can be bought.

Although the parking assist system is small, it is a very complex system within the vehicle. There are many subsystems and components that must cooperate for the system to work as it should. In general, automatic parking systems consist of three main components: the target position designation, path planning and path tracking by active steering [17]. The subsystem includes sensors, cameras, control center, computers, and user-interface, to mention some. There are 3 main sensors used in modern parking system technology:

- Ultrasonic sensor (most common) – Emits soundwaves that humans cannot hear.
  Measures the distance using D =0.5 x Time x Speed of sound.
  The speed of sound varies in different temperatures and humidity; therefore, the car has temperature sensors that can compensate.
- Electromagnetic sensor – Emits radio waves at certain frequency. These reflect back when they hit an object and reach the car at a different frequency. The car is then able to detect the change in frequency and calculate the distance.
- Radar – Located behind the bumper bar at each corner of the car pointing out towards the side. Used for rear cross traffic alert.

Having eyes all around the car has become more common in newer cars as well. Reverse view camera, front camera and wide-angle camera on the wing mirrors are stitched together to produce a top view of the car and surroundings. This is also known as a 360-degree camera or birdsview [2].

The most used ultrasonic sensors only work from 3-5 meters from the car, whereas radar can go further down the road. Reverse Autonomous Emergency Braking (AEB) is usually implemented into the system and activates when there is an obstacle too close that the sensors detect or if the radar detects a moving vehicle or pedestrian [2] [10].

From the moment the parking assist system has been activated, it starts collecting data and calculating the needed distances. It also detects any obstacles via cameras and

 ISBN: 978-1-61208-941-6

sensors as well as taking over the steering system. The cameras and sensors act as transmitters and receivers. These signals go into the computer, which then calculates the distance from objects etc. [8].

III. Problem Context

Over the years, cars have been getting bigger. Also, traffic and the number of vehicles on the road have increased over many years. There is a high demand for parking spaces and parking lots. Parking lots are where one of five motor vehicle accidents happen in the US. Some of the accidents happen between two cars crashing, however most of the accidents involve pedestrians walking in the parking lot. Some of the top causes of parking lot accidents includes drivers entering a parking lot and keeping an eye on a spot, not focusing on the surroundings such as other cars, pedestrians, or obstacles. Another reason is drivers who must back out of their spot, not being able to see all the surroundings. Pedestrians and other moving vehicles are at that point in the driver's blind zone [4].

Statistics from National Centre for Statistics and Analysis (NHTSA) from the year 2007 illustrate numbers of non-traffic fatalities and injuries (Figure 1).

Event	Fatalities	Injuries
Non-occupant in Non-traffic Crash: Backing Vehicle	99	2,000
Non-occupant Struck by Driverless Vehicle	5	<500
Non-occupant in Non-traffic Crash: Forward-Moving Vehicle	106	3,000
Total (approx)	210	5,000

Figure 1. NHTSA Non-traffic fatalities and injuries 2007 [4].

A safe and reliable parking assist system can provide a securer and less stressful parking experience for the driver and help avoid accidents and damages to other cars, obstacles, and people nearby. Sensor and radar systems provide the driver with 'eyes' surrounding the car that easily will alert or activate emergency brakes in unforeseen cases and human behavior. The test, verification, and validation of all the subsystems and the parking assist system as a whole need to be thoroughly executed to provide a reliable system for the user.

IV. System Decomposition

There are many components and subsystems that have to work together in order for the autonomous/semi-autonomous parking assist system to complete its task. There are various conditions the system must fulfil:

- Detect obstacles in the surrounding environment.
- Measure and estimate the distance to obstacles.
- Provide a planned route to park.
- Provide a real-time display to the driver during the parking sequence [5].

These conditions are done by separate interconnected subsystems, as seen in Figure 2.

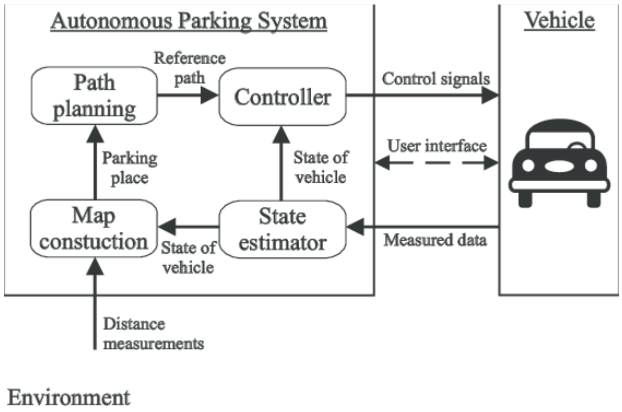


Figure 2. Components of the Parking Assist System [5]

According to Szadeczky-Kardoss and Kiss, the ABS sensors assembled on the vehicle detect the displacement of the wheels on the car. Furthermore, this data can be used to calculate an estimated position and orientation. This estimated state is then used by the mapping and controller modules. For the map to be developed, additional data is also required about the environment. Ultrasonic sensors (most commonly) are used to measure the distance to nearby objects and obstacles. Based on this data, a map can be assembled to near accuracy [5].

Throughout the path planning, a reference path is calculated, which connects the initial and the desired final configurations. In this path planning stage, there are certain constraints that need to be taken into consideration. This can be collision avoidance and the non-holonomic performance described in the model. Ultimately, the tracking control algorithm is used to track the reference path [5].

In Figure 3, the system environment for the parking system is illustrated. The inputs are sensor data that contains information on the vehicle state. For example, vehicle speed, steering position and information from environmental sensors that register objects on the right and left side of the vehicle. In terms of output, the system possesses an interface to the vehicle actors where the vehicle steering angle and velocity will be set [9].

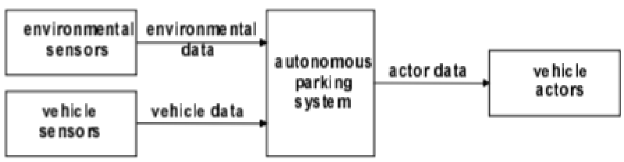


Figure 3. System Environment [9]

V. Test, Verification and Validation

Testing, verifying, and validating the parking assist system is one of the most important steps in the process of product development. This is where the subsystems and the system are tested to see if it fulfils the requirements or if certain things need to be improved. This step can also involve the safety of the product and determines the acceptable level.

A. Parking Assist Test Company (VBOX Racelogic)

VBOX Racelogic is a company that provides equipment for the manufacturers to make sure the parking assist system can be easily tested and validated. Using a VBOX 3i datalogger, RTK base station and survey trolley a parking space can be marked out and plotted to more than 2 cm accuracy. This makes an exact resemblance of the space which then can be uploaded to the VBOX test suite (Software) where the parking assist test can be configured and run. Another feature is that the user can set defined pass and fail conditions in the system to fit the user conditions.

The parking space only needs to be mapped out once, meaning the test can be run multiple times in the same space and variations can be minimized more easily [6].

When the test is running, real-time data can be seen on the screen, as shown in Figure 4. It illustrates the cars outline, the closest point of contact, as well as the pre-marked parking space. This gives the user an instant insight to the pass and fail status of each test, saving valuable testing time.

Figure 4. VBOX Test Suite Park Assist [6]

B. *BMW X5 Parking Assist Test*

The parking assist technology is one of the best technologies within parking features existing today. The technology has been around since 2015 and BMW, Mercedes-Benz, Tesla and Volvo have all introduced autonomous parking assist [1].

Looking into BMW's parking technology, a test on the parking features for a BMW X5 xDrive30d was performed by Paul Maric [2] where he went through all the features and put the system to test in an empty parking lot, using suitcases, himself and two other cars as obstacles. The two cars made it possible to create different scenarios like parallel parking, perpendicular and remote parking.

During these tests, the system is tested in its entirety, not each subsystem separately. When testing the perpendicular parking, the two extra cars were parked within the lines leaving one parking spot in between them. The car then drove slowly past the empty spot to estimate the size. After passing the spot, the car will either alert the driver that the spot was too small, or big enough, so the parking operation can commence. In this case, the spot was big enough. If the driver accepts, the car will automatically set itself to the reverse or drive and park within the lines. At this point, all the sensors are active and the driver should not do anything apart from hovering over the break as an emergency if something were to happen. The driver also gets a 360-degree view of the car from above, also known as birdsview, making it easy to see how the parking is going and that the car will be parked within the lines. Once the parking is complete, the car will alert the driver on the screen and the assisted car parking is complete.

Very similarly to the perpendicular parking, the parallel parking begins by measuring up the space in between two vehicles and the car will alert the driver if the space is suitable. The driver can then accept and start the parking sequence where the car then automatically will set itself into the needed gear and make its way into the space without hitting any of the other vehicles. If something was to interrupt the parking sequence whilst its running, for example a pedestrian walking behind the car, the parking will be paused and resumed when the obstacle is outside of the path.

As Paul Maric mentions, the exit assist may be needed in many cases when the parallel park feature has been used to get the vehicle into the spot. Similarly, starting the other parking assists, one selects exit assist, and it will have the driver confirm the direction of exit. The system will the reverse the car to clear room for the car in front, then drive out of the spot until the driver can take over [2].

Individual tests for the subsystems can also be performed. Regarding a senor test on a vehicle, a simple multimeter test can be used. Other practical tests can also be performed easily by putting the sensors to the test. Normally, objects are placed in certain locations then a scanning tool is used to see if the object is being correctly detected [7]. Placing an obstacle in front or behind the car and driving towards it will show if the sensors pick up on the obstacle in time and how the sensors beep at different distances.

Understanding how the system works can reduce the diagnostic times. Also, knowing what part of the system to test will help prevent installing parts that do not fix the problem and make the repair or improvement cheaper and quicker [3].

VI. DISCUSSION

Systems thinking gives the industry a new opportunity to see the bigger picture and seeing the system with all perspectives and stakeholders involved. When looking at the parking assist system, the main goal for this system is to park without hitting anything and park in a sufficiently large spot. For the system to work and reach its goal, we need reliability and accuracy.

The goal seems like a simple case to solve, and that may be if all the impacting factors are at the perfect state. This is, however, very rare and unlikely in a real-life situation. What makes the whole situation so complex is the randomness and human behavior that occurs. A case scenario could be driving on one of the busiest streets in Oslo wanting to parallel park along the street. In this case, we do not have an empty parking lot where nothing can go wrong. We have randomness and human behavior. There are cars everywhere, trams and tram tracks, buses and people and pets walking all over the place. This means there are many unforeseen situations that can occur and effect the parking assistance system. What happens if a person walks behind the car when parking? What happens if a car behind drives too close during an active parking sequence?

What would be some of the potential test criteria or requirements to have a successful and reliable product?

- Perform parallel and reverse parking without hitting anything.
- Measure and determine a suitable parking spot that will fit the car and allow room to exit and enter the car.

- Park within a time limit of 30 seconds (for example) from accepting and activating the parking.
- Emergency breaks in case of unforeseen behaviour.
- Perform parking assist on the exit of the parking spot.
- Provide clear instructions to the driver during the parking procedure.

Examining the types of components in the system, in order to fulfil these criteria, the sensors may not need to be the absolute best and most expensive to complete the same tasks. To perform a parking operation, all the subsystems must cooperate to achieve the goal and if the components are good enough to fulfil the set requirements with a reasonable margin, it should be acceptable.

VII. Systems Thinking Approach

A. CATWOE from company perspective

The CATWOE approach illustrates the different actors and stakeholders and their perception on a topic. In this case, the CATWOE approach is done from the company's perspective. This approach is from the systems thinking framework [16]. Perspectives are representations of an individual's truths based on their knowledge of the world, and we use multiple forms to contextualize and communicate these perspectives (i.e. verbal, written, graphical) [16]. Using the CATWOE approach can give an overview of what to focus on and perhaps a clearer understanding when making certain decisions and what to prioritize. An example could be making a user-interface that is so complicated that only software engineers would understand it. The user-interface could be very good; however, it is then important to remember who the customers are, who in this case are the drivers. The assumption should be that the divers are not software engineers, and the user-interface needs to be user friendly and easy to understand and comprehend for anyone.

Table I. CATWOE Company Perspective

ASPECT	DESCRIPTION
CUSTOMERS – who are the beneficiaries/victims?	Drivers.
ACTOR(S) – who are the implementers?	The car manufacturers (software, electronics, and test department).
TRANSFORMATION – what does the system do? What are the inputs and what transformation do they go through to become the output?	Signals and communication between different subsystems.
WORLD VIEW – what point of view justifies its existence to the customers? What point of view makes this system meaningful? The big picture and its impact.	Verify the subsystem. Validation of system that it works as it should.
OWNER – who has the authority to change the system?	Authorities like DMV (Department of Motor Vehicle in USA), DVLA (Driver and Vehicle Licensing Agency in England) or Biltilsynet in Norway can set requirements.
ENVIRONMENT – What are the external constraints?	Guidelines, rules, and regulation in the traffic. Randomness and human behavior.

Customer - Drivers would be the main customers of this product and will be most beneficial for them as they are the direct users. The general system technology would, however, be technology that car manufacturers would want to get and interoperate into their line of products to improve the overall performance and possibilities, and also keep up with the technological development.

Actors - The car manufacturers are the ones responsible for implementing the system and making sure the system works and fulfils certain criteria before being released into the final product stage. Within the car manufacturer structure, there are many departments, meaning the software and electrical department mostly will oversee implementing and assembling the system together. In addition, they will be working with the test department to perform tests that verify and validate the different subsystems, then also improve the system if needed. The Department of Motor Vehicles, or Biltilsynet in Norway, are not implementers but may, however, set requirements as to how safe the system must be.

Transformation - The main transformation for this system will be signals. The whole system is made up of sensors, radars, modules that calculate distance and a computer. Simply broken down, one example can be the sensor measuring the distance to the car behind it. This data is sent to the computer which then sends this to the brake system. The brake system can then perform an action of breaking when the car reaches a certain point.

World View – The world view aspect investigates what point of view justifies its existence to the customers and the meaningfulness of the system. For this case, verifying the subsystem and validating the car assist system and making sure it works as it should and fulfils the requirements is what justifies its existence to the customer.

Owner – The owners of the system or the authorities that have the power to influence or change the system in the case of car parking assistance is the DMV (Department of Motor Vehicles in USA), DVLA (Driver and Vehicle Licensing Agency in England) or Biltilsynet in Norway, to mention a few from different countries. They have the authority to make changes or set requirements and constraints to the system, if needed.

Environment – some of the external constraints for the system are guidelines, rules, and regulation in the traffic. Another external constraint can be randomness and human behavior.

B. Systemigram

The term "Systemigram" is derived from "Systemic Diagrams," and has been used to "bring context to the meaning of togetherness" [11] [13] [14]. A system diagram or Systemigram will be used to map the parking assist system. This systems thinking tool is used to explain the interactions between several, interrelated elements and is a great tool to get an overview over all the elements and decompose complex systems [12]. The Systemigram was developed by Boardman and provides a powerful tool for the analysis of systems first described in written form [15]. In this diagram, the mainstay is displayed with the dark and light green bubbles and the path lines are thicker than the rest. The mainstay shows the main steps in the process and shows how it goes from a PAS test to a verified PAS product. See Figure 5.

 ISBN: 978-1-61208-941-6

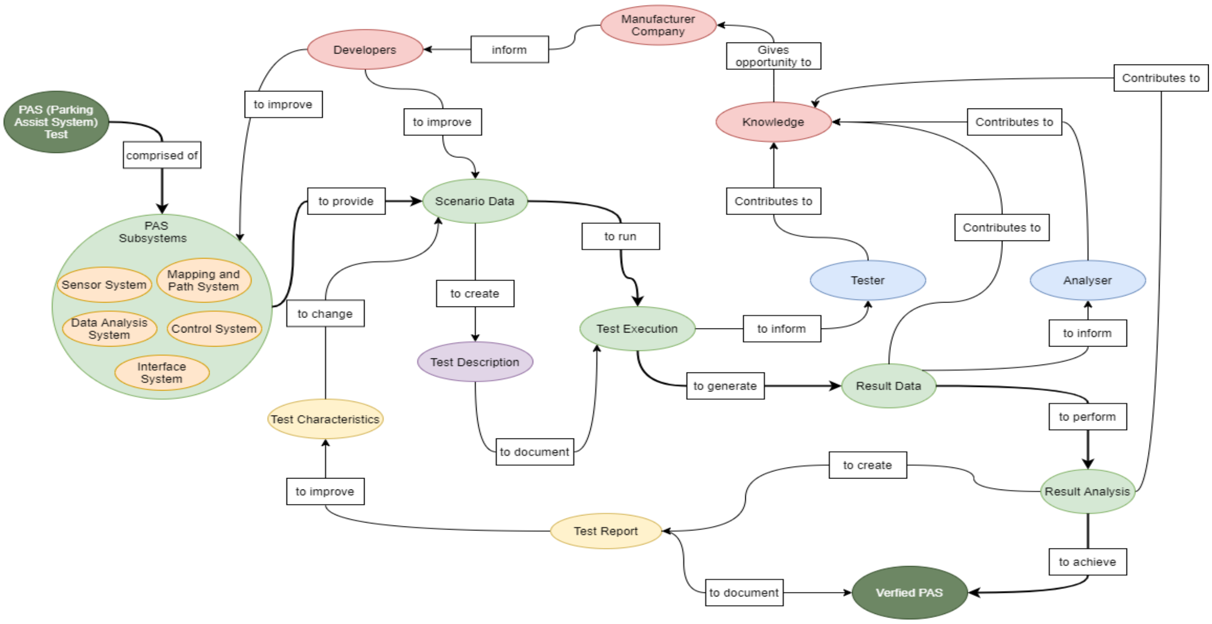


Figure 5. Systemigram of the PAS

VIII. CONCLUSION

The PAS technology is, as mentioned, a small system. However, car manufacturers have prioritized this technology and kept developing it together with the overall vehicle, but why? As explained in the problem context, a large percentage of accidents happen in parking lots and parking garages, not only to other vehicles, but people. Developing a smart parking assist systems in vehicles can help reduce the number of accidents, if not avoid them all together.

The tests and verification of these systems are done thoroughly and there are even companies that produce specific equipment and software for the testers or users to get a good overview and insight in every stage of the parking process. The new modern-day cars are full of parking sensors, radars and, in most cases, also equipped with 360-degree cameras and much more technology and features, such as Reverse AEB (Autonomous Emergency Braking), reverse assist, remote parking etc. Having all these features activated would reduce the risk of accidents as the car would pick up on everything the driver does not see, and even emergency brake, if needed.

The parking assist system technology has been and will be around for a long time. Looking at how much this small and complex technology has evolved since 2003 when the first rear parking sensors were introduced, it is safe to say it will not stop at semi- or autonomous parking assist and will be evolving into technology and new smart solutions we have not even thought of yet.

The purpose of this paper is to propose a new way of looking at a complex system. Using systems thinking approaches as CATWOE and Systemigram, we can look at the system with more perspective and with the involved stakeholders. Although each component and subsystem are important, they can give the system a new view when looking at it with a wider perspective and seeing the whole system.

REFERENCES

[1] S. Gautam, "A Brief History of Car Parking Technology - Get My Parking Blog.", 2019. [online] Get My Parking Blog. Available at: <https://blog.getmyparking.com/2019/04/08/a-brief-history-of-car-parking-technology/> [Accessed 30 June 2021].

[2] P. Maric, "How do parking sensors work? Parking tech, radar & remote hands free explained!", 2020. [online] Youtube.com. Available at: <https://www.youtube.com/watch?v=uDpaB_1kIdA> [Accessed 9 July 2021].

[3] Rts.i-car.com, "Understanding The Park Assist System.", 2017. [online] Available at: <https://rts.i-car.com/collision-repair-news/understanding-the-park-assist-system.html> [Accessed 9 July 2021].

[4] D. Moore, "Parking lot accidents: statistics, causes, and liability - MyParkingSign Blog.", 2021. [online] MyParkingSign Blog. Available at: <https://www.myparkingsign.com/blog/parking-lot-accidents/> [Accessed 6 July 2021].

[5] E. Szadeczky-Kardoss and B. Kiss, "Path Planning and Tracking Control for an Automatic Parking Assist System", 2008. [online] Available at: <https://www.researchgate.net/publication/225220433_Path_Planning_and_Tracking_Control_for_an_Automatic_Parking_Assist_System> [Accessed 6 July 2021].

[6] Zen Micro Systems. 2020. "Park Assist - Zen Micro Systems. " [online] Available at: <https://www.zenmicrosystems.co.in/park-assist/> [Accessed 10 August 2021].

[7] Rts.i-car.com. 2016. "Typical Calibration Requirements For Park Assist Sensors. " [online] Available at: <https://rts.i-car.com/collision-repair-news/typical-calibration-requirements-for-park-assist-sensors.html> [Accessed 10 August 2021].

[8] S. Gautam, "Smart Parking Assist Explained: How it Works - Get My Parking Blog. ", 2020. [online] Get My Parking Blog. Available at: <https://blog.getmyparking.com/2020/06/26/smart-parking-assist-explained-how-it-works/> [Accessed 12 August 2021].

[9] O. Buhler and J. Wegner, "Automatic Testing of an Autonomous Parking System using Evolutionary Computation. SAE Technical Papers," 2004. [online] Available at:

<https://www.researchgate.net/publication/267402961_Automatic_Testin_of_an_Autonomous_Parking_System_Using_Evolutionary_Computation> [Accessed 9 July 2021].

[10] G. Voinea, C. Postelnicu, M. Duguleana, G. Mogan, and R. Socianu,, "Driving Performance and Technology Acceptance Evaluation in Real Traffic of a Smartphone-Based Driver Assistance System. International Journal of Environmental Research and Public Health", vol. 17, no. 19, p.7098, 2020.

[11] R. Haugen and M. Mansouri, "Applying Systems Thinking to Frame and Explore a Test System for Product Verification; a Case Study in Large Defence Projects. " INCOSE International Symposium, vol. 30, no. 1, pp.78-93, 2020.

[12] C. Witte and M. Mansouri, "Analyzing the effects of connecting Norway's remote communities to large cities using the Systems Thinking approach," INCOSE International Symposium, vol. 30, no. 1, pp. 1219-1234, 2020.

[13] J. Boardman and B. Sauser, "Systems Thinking: Coping with 21st Centurty Problems." 2008, Boca Raton, FL: Taylor & Francis / CRC Press.

[14] M. Mansouri, B, Sauser, and D. Boardman. "Applications of Systems Thinking for Resilience Study in Maritime Transportation System of Systems.", 2009.

[15] R. Edson, "Systems Thinking. Applied. A Primer. ", 1st ed. ASysT, p.21, 2008.

[16] B. Sauser, M. Mansouri and M. Omer. "Using Systemigrams in Problem Definition: A Case Study in Maritime Resilience for Homeland Security. Journal of Homeland Security and Emergency Management", vol. 8, no. 1, 2011.

[17] H. Jung. "Semi - automatic parking slot marking recognition for intelligent parking assist systems. The Journal of Engineering", no. 1, pp.8-15, 2014.

CROWD_SZ: A Large-scale Multi-view Crowd Counting Semantic Dataset

Jiajia Wu
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
20194246023@stu.suda.edu.cn

Yufeng Lin
Public Security Bureau Command Center
Suzhou Industrial Park
Suzhou, Jiangsu, China
18913585632@163.com

Cheng Wu
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
cwu@suda.edu.cn

Jin Zhang
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
zhangjin1983@suda.edu.cn

Lijun Zhang
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
zhanglijun@suda.edu.cn

***Abstract*—In recent years, there have been many large social gatherings and stampedes. High-density crowd counting and density estimation have become a research hotspot in the field of video surveillance. However, traditional datasets expose the limitations of a single perspective and limited crowd size, which cannot meet the research needs of a wide-area place. This paper proposes a new large-scale multi-view dataset, taking the urban life square near Jinji Lake in Suzhou city, Jiangsu Province, China as the research object. A single camera cannot cover the whole place, so we collect surveillance images from multiple perspectives. The low-altitude monitoring image has obvious human characteristics, while the high-altitude monitoring image provides the trend of crowd distribution. Combining these two kinds of information, the trend of crowd change can be predicted more accurately. This dataset is characterized by rapid crowd change in a short time, large aggregation scale and complex illumination conditions, which brings new challenges to crowd counting research.**

Index Terms*—*Crowd counting; Semantic understanding; Data fusion.

I. Introduction

With the rapid development of urbanization, more large-scale competitions, cultural exchanges, and entertainment activities are held. The actual crowd is often greater than the number of people that can be accommodated in the venue, which has caused a series of unexpected safety problems [1]. To improve event management and safety, related research has shifted focus to the field of crowd technology [2]. Thanks to the widespread use of video surveillance systems, the all-around installation of video capture equipment provides more research data for the field of crowd density estimation, making it possible to accurately count crowds in dense places. Different from other datasets, the pedestrian features in the crowd dataset are small and fuzzy, which makes it more difficult to capture. In addition, changes of perspective, over-dark or over-exposed environmental illumination, and crowd occlusion hurt feature extraction, as shown in Figure 1.

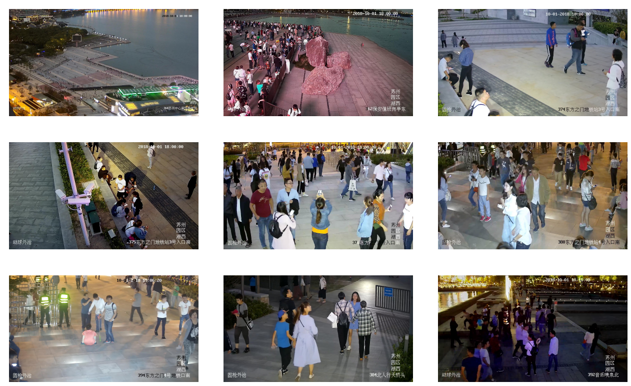


Figure 1. Example from CROWD_SZ.

At present, a lot of research is devoted to crowd counting in natural scenes. Some public places such as squares and stations also have a strong demand for accurate real-time crowd counting [3]. Traditional crowd counting methods are mainly divided into two categories: detection-based methods [4]–[7] and regression-based methods [8]–[10], [21]. Early crowd studies mainly used detection-based methods. First, assuming that the crowd is composed of individuals, use the sliding window detector to detect the crowd and calculate the crowd count. In a dense crowd environment, however, detection-based methods are difficult to solve the problem of serious occlusion inside crowds. Therefore, regression-based methods were introduced. By learning the mapping of a feature to the number of people, these methods extract the foreground features and then use a regression function to estimate the number of people. The mapping process usually adopts the linear [12] or piecewise linear [11] function model. In recent years, more attention is paid to the crowd counting technology based on deep learning. It is different from the traditional crowd counting method, which uses multiple convolutional neural networks to extract head features of different sizes. This method can get better prediction results for densely distributed regions. Zhang et al. [13] proposed a crowd counting model (Crowd CNN) based on a six-layer convolutional neural network. By alternately optimizing the true value of the crowd count and the true density map of the crowd, better robustness was obtained. Bai et al. [28] extracted the global coarse-grained motion of the crowd from the perspective of high altitude. The local fine-grained density characteristics of people with line of sight occlusion are extracted from the low altitude perspective, and a holographic model of the temporal and spatial evolution of crowd situation is established. Fu et al. [14] designed two ConvNet cascaded classifiers to estimate crowd density by optimizing the convolutional neural network. Zhang et al. [15] proposed a Multi-Column Convolutional Neural Network (MCNN) structure to extract multi-scale image features. After image inputting, the crowd density estimation map was obtained. Finally, the density map was integrated to obtain the estimated crowd count value. Hu et al. [16] used the ConvNet structure to extract crowd features and then used two signals of crowd count and crowd density to learn crowd characteristics and estimate the number of people. Jiang et al. [17] focused on the improvement of density maps, using additional multi-scale markers to increase the diversity of deep neural networks, and achieve the high-performance crowd density map estimation. From the traditional detection-based and regression-based methods to the application of deep learning frameworks, most of the current methods count by extracting human head features of different scales. Unfortunately, such methods cannot effectively deal with the problem of target detection in dense crowds in natural scenes. The main problems are:

- The change of perspective and the position of equipment will lead to a great change in the size of the human head in dense crowds, and then the feature extraction method based on a single or finite-size convolution kernel has difficulty extracting the full size of a human head.
- The number of targets contained in a dense crowd varies greatly, and the number changes significantly in a short time. An image can contain several thousands of people. Therefore, the method based on multi-frame fusion detection can not be applied to the estimation of the number of people with obvious differences.
- Research objectives in a dense crowd are often unevenly and irregularly distributed. It is difficult to describe the change in the global situation using methods based on statistical reasoning.

In view of the typical weaknesses of the dense crowd target detection methods, existing studies lack of dense crowd scenarios to meet the research needs. Some existing crowd datasets are mostly ideal, with a single scene and unchanged illumination, which cannot fully reflect the complexity of the problem, such as UCSD [18], Mall [19], WorldExpo'10 [20], UCF_CC_50 [13], Shanghaitech [15] and so on. UCSD [18] is the first dataset in the crowd counting field. It consists of 2,000 frames of images and pedestrian annotations in each frame, and the video frames are extracted from a single scene. The density of images in this dataset is low, with an average of 15 people in each image. Chen et al. [19] collected a new dataset Mall with different illumination conditions and crowd density. The images in this dataset have a higher crowd density. But, like UCSD, they are all part of a single video sequence, so the scene does not change. WorldExpo'10 [20] contains 108 videos in 5 different scenes, with a total of 3,980 frames of images. UCSD and Worldexpo'10 contain only low and medium density scenarios and lack high-density scenarios. The UCF_CC_50 dataset [13] specializes in collecting ultra-highdensity crowd scenes and contains only 50 images. The generalization ability of the training network with a limited number of training samples is reduced, which affects the test results. The Shanghaitech dataset [15] has better diversity than previous datasets in terms of scenes and density levels. It is divided into two parts: part_A (including images of a high-density crowd) and part_B (including images of a lowdensity crowd). It contains 1,198 images with 330,165 annotations. Although these datasets provide some images for counting, they are lacking in sample number, image complexity, and scene diversity. Qi Wang et al. [22] proposed a new dataset Nwpucrowd, with 5,109 images and 2,133,238 annotations, which has been greatly improved in terms of quantity and provides a platform for researchers to compare the calculation results of the test set. Sindagi et al. [23] proposed a dataset called JHU-CROWD++, which was collected under different scenes and environmental conditions, including some images based on severe weather and illumination changes. However, these two datasets did not fully consider background interference factors in the nature scene, and the identification of human body contour is very unsatisfactory in the high-altitude image with an ultra density of the crowd.

This is the motive for our work. In response to the existing problems of the above datasets, this article introduces a large-scale multi-view crowd counting dataset. Table 1 describes the parameter comparison between our dataset and other typical crowd datasets. We collected 5,610 images and 1,738 video files from the monitoring equipment. The images come from different heights and angles (including two high perspectives and nine low perspectives), and the illumination of the scene changes significantly at night. In terms of perspective selection, to facilitate the research of multi-perspective fusion algorithm, the dataset fully considers the basic principle that a high-altitude perspective must include a low-altitude perspective coverage area, which mainly reflects the interrelation

TABLE I. COMPARISON BETWEEN CROWD_SZ AND OTHER CROWD DATASETS.

Dataset		Resolution	Images	Min	Max	Multi-view	Density change
Shanghaitech	Part_A	different	482	33	3,139	✓	✓
	Part_B	768*1,024	716	9	578		
UCF_CC_50		different	50	94	4,543	✓	✗
UCSD		158*238	2,000	11	46	✗	✗
Mall		480*640	2,000	13	53	✗	✓
WorldExpo_10		576*720	3,980	1	253	✓	✓
CROWD_SZ		different	5,610	1	673	✓	✓

 ISBN: 978-1-61208-941-6

(a) Bounding box annotation (b) Human head annotation

Figure 2. Annotation.

and spatial complementarity of different perspectives. At the same time, the images selected from the dataset also highlight the different crowd density levels, illumination conditions, perspective distortion, and other conditions.

In summary, the main contributions of this article are as follows:

- We propose a new large-scale multi-view dataset for crowd counting and density estimation. The dataset includes 5,610 images and 1,738 videos, which makes up for the lack of diversity in traditional datasets;
- For the stock data of the dataset, a large number of complete head annotation files are carefully prepared;
- For the key elements of dense crowd vision, different scene video frames with good statistical dispersion are provided;
- Fully considering the basic principle of ”the high-altitude perspective must include the low-altitude perspective coverage area”, our dataset provides global and local images, and allows to calculate the number of people in wide-area places according to the time correlation and spatial complementarity between them.

The rest of the paper is structured as follows. Section II describes the annotation method and the classification of the dataset. Section III gives the specific nature of the dataset and some statistical information of the data. Section IV provides the experimental procedures and data of the two methods, and performs crossdataset verification to verify the generalization ability of the dataset. Section V concludes the paper.

II. Annotation

A. Dataset

In the crowd dataset, image acquisition equipment and acquisition scenarios are the main reasons for the deviation of the dataset. To eliminate the deviation, we collected videos and images in the CROWD_SZ dataset from monitoring equipment at different locations in Suzhou Jinji Lake Fountain Square. We recorded the specific location of each perspective and the specific time of each image to ensure that the image time between different perspectives is consistent with the subject of observation. For video files with different perspectives, we save a one-minute video as one file, which also guarantees consistency between different perspectives.

B. Classification

CROWD_SZ is divided into image sets and video sets. The image set contains high-altitude images and low-altitude images, with high-altitude images having two perspectives. The low-altitude images are divided into nine perspectives. Each folder contains 510 images and the corresponding pedestrian header annotation file. Video sets and image sets follow the same classification criteria; Each subclass contains 158 video files and each subclass contains a 1-minute video. We divided the annotation files into training set and test set according to a 3:1 ratio.

C. Annotation method

In the direction of pedestrian detection, many datasets use bounding box annotation [24] and pedestrian torso line annotation [25]. In our dataset, there is a large number of images with pedestrian occlusion and overlap. If the above two annotation methods are still used, multiple boundary boxes and pedestrian trunk lines will overlap extensively, and the pedestrian in the image cannot be accurately detected, resulting in error counting results. Considering the counting requirements, we annotate the head of the pedestrian in the image. As shown in Figure 2, the head position of each pedestrian is marked with a red cross. The marking process is mainly divided into two parts: crowd image labeling and transforming crowd image labeling into a crowd density map.

The label density map generation process is as follows: First, x_i represents the center coordinates of the human head. If there is a human head at a specific position x_a, it can be expressed as $\delta(x - x_a)$, which means that there is a actual person at the x_a coordinate position. If there are N heads in a picture, then this picture can be represented as follows:

$$H(x) = \sum_{i=1}^{N} \delta(x - x_i) \quad (1)$$

The density map of the image is generated by the function and gaussian kernel convolution. Due to perspective distortion in the scene, the size of each human head needs to be considered to determine the diffusion function before generating the density map. The size of the head is usually the distance between the centers of two adjacent people. Therefore, we adopt an adaptive method to determine the parameters of each person. The formula for generating the final density map is as shown in (2), where G represents the Gaussian kernel, σ is the standard deviation of the Gaussian kernel, and β is a set value, usually 0.3. Suppose there are k people around this person. d represents the average of the sum of the Euclidean distances of the head from its k neighboring heads.

$$F(x) = \sum_{i=1}^{N} \delta(x - x_i\ ) * G_{\sigma_i}\ (x), with\ \sigma_i = \beta \overline{d_i} \quad (2)$$

III. The nature of CROWD_SZ

A. Image collection location and pixel size

Compared with other datasets, the original sizes of the images in our dataset are $1,920 * 1,080$, $2,560 * 1,440$, etc., while the image sizes of other existing datasets mostly do not exceed $1,000 * 1,000$. For example, the image size in UCSD is $158 * 238$. The picture size in Mall is $480 * 640$.

 ISBN: 978-1-61208-941-6

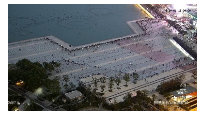
(a) Low density

(b) Medium density

(c) High density

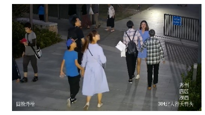
(d) Ultra high density

Figure 3. Examples of crowd images with different density levels.



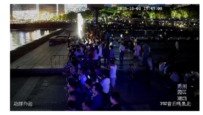
(a) High altitude image

(b) Strong illumination image

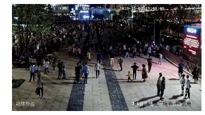


(c) Close up scene

(d) Low altitude image

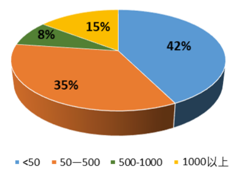


(e) Weak illumination image

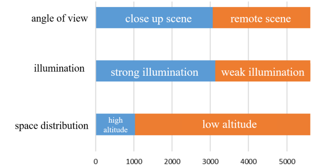


(f) Remote scene

Figure 4. Examples of different types of images in CROWD_SZ.



(a)



(b)

Figure 5. The characteristics of CROWD_SZ.

B. Density distribution

In terms of density (here, we define density as the number of people contained in a single image.), as shown in Figure 3, we divide it into four density levels:

- Low density: the count value is between zero and fifty.
- Medium density: the count value is between fifty and five hundred.
- High density: the count value is between five hundred and one thousand.
- Ultra high density: count value above one thousand.

Compared with other datasets with a single density distribution, our dataset has made progress in the diversity of density. We counted all the pictures included in the dataset, and finally got the distribution of people as shown in Table II. In this dataset, low-density images account for a relatively high proportion of 42%, medium-density images account for 35%, high-density images account for 15%, and ultra high-density images account for 8%, which is enough to meet the research of different density images demand.

C. Diversity

Scene diversity is an important attribute of the dataset. Our dataset contains images taken by cameras at different heights and angles. The scenes, illumination, and pedestrian forms are diverse. As shown in Figure 4, the images in the dataset can be divided into a high-altitude image and low-altitude image according to spatial distribution, strong illumination image, and weak illumination image according to illumination conditions, and close-up scene image and remote scene image according to perspective. Figure 5 depicts the statistical data of the above-mentioned scene diversity. In terms of spatial distribution, high-altitude image data accounts for 1/5 of all data, while low-altitude image data accounts for 4/5 of all data. In the illumination condition category, the number of images with strong illumination accounts for 1/2 of the total, while the other 1/2 are images with weak illumination. Similarly, under the perspective position category, the number of close-up images and remote images each account for 1/2 of the total.

TABLE II. DIFFERENT DENSITY DISTRIBUTION IN CROWD_SZ.

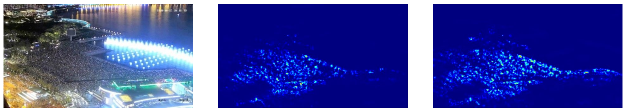
Density	Low	Medium	High	Ultra high	Total
No. of images	2,389	1,941	460	820	5,610

IV. EVALUATION

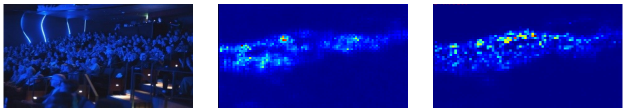
In this section, we evaluate some typical crowd counting algorithms on the CROWD_SZ dataset. We choose MCNN [15] and CSRNet [26] as the benchmark algorithms for low-altitude image processing in the dataset. Here, MCNN uses three columns of different scale convolution kernels to adapt to different scales of human head sizes, and finally combines the three columns of neural networks to obtain a density map. The CSRNet network model is divided into a front-end network and a back-end network. The front-end network will use VGG [27] (Visual Geometry Group Network) with the fully connected layer removed, and the back-end network will use a hollow convolutional neural network. Its purpose is to generate high-quality crowd density maps while maintaining the resolution while expanding the perceptual field.

A. Image preprocessing

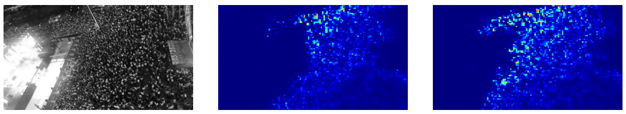
One of the advantages of our dataset is that it contains many night scenes. Take the challenging scenes shown in Figure 6 as an example. To recognize the pedestrian characteristics better, we preprocessed the images, the image needs to be grayed first to obtain the density map. This preprocessing



(a) (b) (c)



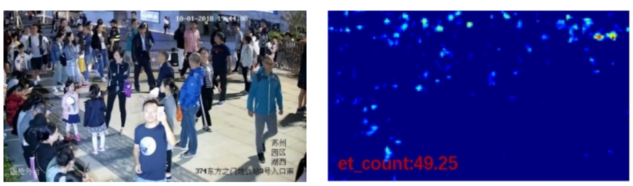
(d) (e) (f)



(g) (h) (i)

Figure 6. Image preprocessing: (a)(d)(g) is the original image, (b)(e)(h) is the density map obtained from the input original image, and (c)(f)(I) is the density map obtained by graying the original image.

(a) Input (b) Output



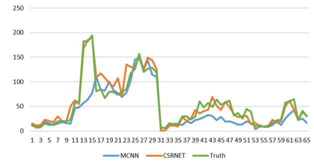
(c) Input (d) Output

Figure 7. Input and output.

step helps to identify the crowd more comprehensively and carefully, thus making the numerical calculation more accurate. We also selected images from other datasets at night or in week-illumination for verification. Experimental results show that grayscale pretreatment can improve the accuracy of recognition.

B. Experimental analysis

To evaluate two crowd counting algorithms based on the CROWD_SZ dataset, we randomly selected 65 images in the dataset, including 31 images with relatively sparse scenes and 34 images with crowded scenes. Two algorithms, MCNN and CSRNet, are used to estimate the crowd density. The experimental results are shown in Figure 7. We select a low-altitude image as the input data to get the specific number of people and the output density map.



(a)

(b)

Figure 8. (a) is the comparison of prediction results of MCNN and CSRNet. (b) is the installation location distribution and corresponding resolution of monitoring devices.

As shown in Figure 8, we list the estimated number of people in the same image obtained by two algorithms, both of which predict the image with higher accuracy. The detailed description is shown in Table III, we use MAE (Mean Absolute Error) as evaluation criteria, the smaller the value, the higher the estimation accuracy. In crowded scenarios, CSRNet has a higher estimation accuracy. On the contrary, MCNN has a higher estimation accuracy in sparse scenes. Select the high-altitude density map and the low-altitude density map of the same time. Finally, as shown in Figure 9, comparing these two density maps, we can find that the local density is consistent with the high-altitude density, which also shows that our count is more accurate and reflects the current crowd density.

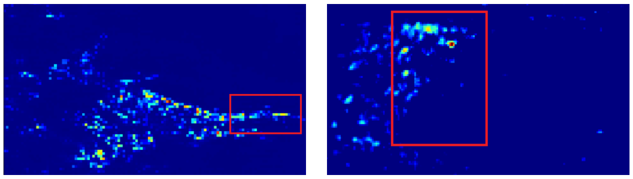
At the same tine, we choose the JHUCROWD++ for cross-dataset induction. Compared with other crowd datasets, JHU-CROWD++ has a large number of images collected under different environmental conditions and scenes, including severe weather changes and illumination changes, which makes it very challenging. We randomly select ten images and evaluate the accuracy of their count. The results are shown in Table IV. We used MCNN and CSRNet to estimate the number of people in the image. The estimated results are the same, which also shows that our dataset has good applicability. However, there are also significant differences between the estimates of MCNN and CSRNet. This shows that CROWD_SZ is more challenging.

TABLE III. COMPARISON OF MAE VALUES BETWEEN MCNN AND CSRNET IN DIFFERENT SCENARIOS.

Scene \ MAE	MCNN	CSRNet
Sparse scene	0.02839	0.1429
Crowded scene	0.2979	0.1415

TABLE IV. COMPARISON OF MCNN AND CSRNET'S ESTIMATED COUNT OF IMAGES IN TWO DATASETS.

	CROWD_SZ			JHU-CROWD++		
Number	Truth	MCNN	CSRNet	Truth	MCNN	CSRNet
1	7	9.48	11.42	84	82.93	98.65
2	21	17.47	30.39	41	55.22	53.26
3	408	91.08	246.75	13	11.19	14.04
4	39	39.02	32.87	68	28.17	79.23
5	44	39.6	39.71	69	50.32	120.05
6	65	71.63	67.98	42	37.56	58.18
7	17	15.02	10.59	7	5.91	13.02
8	53	46.24	62.52	188	153.5	272.98
9	13	11.48	16.9	65	66.5	81.98
10	23	15.32	24.27	50	47.62	81.52



(a) High altitude image (b) Low altitude image

Figure 9. Density contrast map of the same area in high altitude and low altitude.

V. Conclusion and Future Work

In this work, we introduce a new challenging crowd dataset for large crowd gatherings. The dataset consists of 5,610 images and 1,738 videos, including a large number of night scenes. This dataset contains more images than the existing crowd dataset and has good temporal and spatial correlations between images under different spatial positions, illumination changes, and perspective changes. In addition, we provide a complete header annotation markup file. In this paper, by using a typical crowd counting target detection algorithm and cross-dataset validation process, it is proved that the dataset is larger, more diverse, and more challenging, which is suitable for practical application and can be used as the basis of dense crowd counting research dataset. In the future, we will use spatial complementarity between high altitude images and low altitude images to count vast areas.

VI. Aknowledgement

This work is a part of the project "Design and consulting service scheme for safety control system of crowd flow in large urban complex and crowd concentrated place". Thanks to Cheng wu, the corresponding author, for giving me the opportunity to participate in the meeting.

References

[1] L. Soomaroo and V. Murray, "Disasters at Mass Gatherings: Lessons from History," Plos Currents, 2012, pp. 1-10.
[2] J. Shao, K. Kang, C. C. Loy and X. Wang, "Deeply learned attributes for crowded scene understanding," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 4657-4666.
[3] E. Stromatias, "Developing a supervised training algorithm for limited precision feed-forward spiking neural networks," Computer Science, 2011, pp. 1-107.
[4] M. Li, Z. Zhang, K. Huang and T. Tan, "Estimating the number of people in crowded scenes by MID based foreground segmentation and head-shoulder detection," 2008 19th International Conference on Pattern Recognition, 2008, pp. 1-4.
[5] L. Wang, L. Xu and M. H. Yang, "Pedestrian detection in crowded scenes via scale and occlusion analysis," 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 1210-1214.
[6] M. Enzweiler and D. M. Gavrila, "Monocular Pedestrian Detection: Survey and Experiments," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 12, 2009, pp. 2179-2195.
[7] O. Tuzel, F. Porikli and P. Meer, "Human Detection via Classification on Riemannian Manifolds," 2007 IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1-8.
[8] U. Khan and R. Klette, "Logarithmically improved property regression for crowd counting," 2016, pp. 123–135.
[9] A. B. Chan and N. Vasconcelos, "Bayesian Poisson regression for crowd counting," 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 545-551.
[10] K. Chen, C. C. Loy, S. Gong and T. Xiang, "Feature Mining for Localised Crowd Counting," British Machine Vision Conference. 2012, pp. 1-12.
[11] N. Paragios and V. Ramesh, "A MRF-based approach for real-time subway monitoring," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, pp. 1034-1040.
[12] V. Rabaud and S. Belongie, "Counting Crowded Moving Objects," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006, pp. 705-711.
[13] C. Zhang, H. S. Li, X. Wang and X. K.Yang, "Cross-scene crowd counting via deep convolutional neural networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 833-841.
[14] F. Min et al., "Fast crowd density estimation with convolutional neural networks," Engineering Applications of Artificial Intelligence, 2015, pp. 81-88.
[15] Y. Zhang, D. Zhou, S. Chen, S. Gao and Y. Ma, "Single-Image Crowd Counting via Multi-Column Convolutional Neural Network," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 589-597.
[16] Y. Hu, H. Chang, F. Nian, Y. Wang and T. Li, "Dense crowd counting from still images with convolutional neural networks," Journal of Visual Communication and Image Representation, 2016, 530-539.
[17] H. Y. Jiang and W. Jin, "Effective use of convolutional neural networks and diverse deep supervision for better crowd counting," Applied Intelligence , 2019, pp. 2415–2433.
[18] A. B. Chan, Z. S. Liang and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-7.
[19] C. C. Loy, S. Gong and T. Xiang, "From Semi-supervised to Transfer Counting of Crowds," 2013 IEEE International Conference on Computer Vision, 2013, pp. 2256-2263.
[20] H. Idrees, I. Saleemi, C. Seibert and M. Shah, "Multi-source Multi-scale Counting in Extremely Dense Crowd Images," 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 2547-2554.
[21] Y. Yuan, Z. Y. Wu and X. Z. Jiang, "Research on Crowd Counting via Deep Convolution Neural Networks," Group Technology & Production Modernization, 2017, pp. 49-53.
[22] Q. Wang, J. Gao, W. Lin and X. Li, "NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 6, 2021, pp. 2141-2149.
[23] V. Sindagi, R. Yasarla and V. M. M. Patel, "JHU-CROWD++: Large-Scale Crowd Counting Dataset and A Benchmark Method," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, vol. 14, no. 8, pp. 1-17.
[24] W. Lee, J. Na and G. Kim, "Multi-Task Self-Supervised Object Detection via Recycling of Bounding Box Annotations," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4979-4988.
[25] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, 2021, pp. 172-186.
[26] Y. Li, X. Zhang and D. Chen, "CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 1091-1100.
[27] Z. Yu et al., "Convolutional networks with cross-layer neurons for image recognition," Information Sciences, 2018, vol. 433–434, pp. 241-254.
[28] L. Bai, C. Wu, Y. M. Wang and F. Xie. "Multiview-Fusion-Based Crowd Density Estimation Method for Dense Crowd," The Fifteenth International Conference on Systems(ICONS), 2020, pp. 50-55.

RailVID: A Dataset for Rail Environment Semantic

Hao Yuan
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
20205246017@stu.suda.edu.cn

Zhenkun Mei
Suzhou Rail Transit Group Co.Ltd.
Suzhou, Jiangsu, China
meizk@163.com

Yihao Chen
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
20204246013@stu.suda.edu.cn

Weilong Niu
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
wlniu@stu.suda.edu.cn

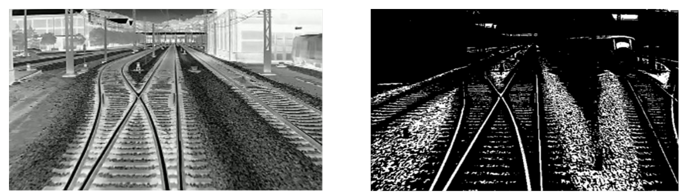
Cheng Wu
School of Rail Transportation
Soochow University
Suzhou, Jiangsu, China
cwu@stu.suda.edu.cn

***Abstract*—At present, rail transit is becoming the main means of urban and intercity fast passenger and freight transportation. Its safe operation is of great significance to protect the life of people, security of property and maintain social stability. Though the degree of intelligent traffic has been improved , there are still many safety risks in the current system by using manual hazard monitoring in the railway. Limited by the particularity and complexity of railway scenes, there are few studies on the perception and understanding of rail transit environment. In this paper, we propose a new rail transit dataset – RailVID. We use the AT615X Infrared thermography from InfiRay to collect data and record different railway scenarios, including carport, depot, and straight. We then propose an improved BiSeNet real-time semantic segmentation network for evaluation. Based on this dataset, we carry out environment perception, environment understanding, and safety decisions on the track area in front of the train, and we propose a solution for fully automatic train operation of rail transit.The dataset we provide compensates for the infrared data that is not in the existing dataset, and our data covers special weather and various conditions. Experiments show that our method achieves a higher Mean Pixel Accuracy in the collected dataset, and the processing speed also meets the real-time requirement.**

Index Terms*—Semantic segmentation; Rail transit; Environmental perception.

I. INTRODUCTION

At present, with the continuous advancement of modernization [1], rail transit has become the main trunk line of public transportation and the main artery of passenger and freight transportation. All trains, however, locate themselves on the rails through a safety control system to ensure the running interval [2], in order to ensure the running safety. In this context, there is a lack of means to perceive and understand the environment. Therefore, when manual observation is limited or signal equipment is faulty, it is difficult to judge the environment ahead, which may lead to serious accidents. In recent years, research on the environmental perception of urban rail transit has become a top priority. At the same time, the successful application of environmental perception in the intelligent assisted driving system of vehicles provides new ideas and solutions to the problems of environmental understanding and environmental perception of urban rail transit.



(a) Orbit image and binarization image

(b) Track lines are more complex than lane lines

Figure 1. Unique problems of track circuit.

In the current period of rapid development of urban rail transit, it has become the mainstream research direction to improve the intelligent level of operation with advanced scientific and technological methods. With the rapid development of deep learning and the progress of sensor technology, the application of deep learning in environment perception is being widely used in rail transit. Its main application fields are turnout fault diagnosis [3], passenger flow prediction [4], carriage security detection [5], train environment perception, etc. Deep learning has great potential in train environment perception. Through the deep learning algorithm processing of the data collected by the video sensor, the driver can improve the warning messages of road conditions ahead, which can effectively reduce the probability of accidents, ensure the safe operation of the train and improve the operation efficiency.

Image semantic segmentation refers to the specific pixel level when recognizing the image, that is, to divide each pixel of the image into a category, which is of great significance

 ISBN: 978-1-61208-941-6

in the environmental perception and understanding of rail transit [15]. However, in the case of poor imaging conditions such as night, rain, fog, and point light interference, semantic segmentation using visible images will lead to a decrease of visual perception. Infrared cameras use thermal radiation to generate thermal images, which are independent of the light source, are less affected by weather, and have longer detection distance. With an infrared camera, images are clearer and more stable at night when there is dark, rain, fog, or point light source pollution [16]. Using infrared thermography for real-time semantic segmentation of track environment can perceive the environment under the condition of poor visible light imaging, enhance the adaptability to special environment and special conditions, and improve the safety and intelligence of fully-automatic train operation. The unique problems of track circuit are shown in Figure 1.

The rest of the paper is structured as follows. In Section II, we present the current development of semantic segmentation and our contributions. In Section III, we introduce the BiseNet network which is suitable for infrared data, and propose an improved BiseNet to enhance the accuracy and meet the real-time demand. In Section IV, we introduce the equipment we used and the data collection scenarios, then we annotate and categorize the data. In Section V, we introduce the allocation and evaluation index of our model. Then, we compare the results of different methods on our dataset and show that our method is better. Finally, we conclude our work in Section VI.

II. Related work

Real-time semantic segmentation technology is mainly used in autonomous driving, which can help vehicles enhance their perception of their surroundings and understand their environment. In 2014, Long et al. [3] proposed a Fully Convolution neural Network (FCN), which replaced the full connection layer in the mainstream classification network with the convolution layer, and upsampled the feature map of the last volume base layer by restoring it to the same size as the input image, so that a prediction can be generated for each pixel. It preserves the spatial information in the original input image, which makes a breakthrough in semantic segmentation. Image semantic segmentation technology based on deep learning has begun to receive wide attention in the academic circle. SegNet [6] introduces the encoder-decoder concept for semantic segmentation. It uses the VGG (Visual Geometry Group) network as the encoder, and the decoder is symmetrically designed. The maximum pooled index at the corresponding encoder layer is called for upsampling. Chen et al. proposed the encoder-decoder with Atrous Separable for Semantic Image Segmentation to improve the Semantic Segmentation accuracy network. The precision of semantic segmentation in autonomous driving environment is further improved by an encoding and decoding structure. Fisher et al. [7] proposed the concept of Dilated Convolution for the loss of location information at the pooling layer, and ensured network accuracy with a larger receptive field. On this basis, Chen et al. proposed DeepLab series [8]–[10], using fully connected Conditional Random field (DenseCRF) and Atrous Spatial Pyramid Pooling (ASPP), respectively. The subsequent DeepLab v3+ [11] proposed a simple and effective decoder module based on the powerful encoder of DeepLab V3. While the DeepLab family improves network performance, it also increases computing costs. The Context Grided Network (CGNet) proposed by Wu et al. [12] uses lightweight framework and context joint feature extraction to improve the real-time performance of semantic segmentation. Lin et al. [17] proposed RefineNet network, which uses residual links to explicitly combine each sub-sampling layer with the following network layer to reduce memory usage and improve feature fusion between modules. Bilateral Semantic segmentation Net (BiSeNet) proposed by Yu et al. [13] adopts a bilateral structure of spatial path and context path, which gives consideration to semantic segmentation accuracy and real-time performance. At the same time, its simple network structure is conducive to later optimization.

Our contribution: Compared with automatic driving semantic segmentation technology of urban roads, the semantic segmentation in visual perception of rail transit has been developed later, and the environment of rail transit is more complex. Moreover, due to the particularity of rail transit, it is impossible to flexibly deploy platforms to achieve large-scale measurement and improvement in public sections. However, the perception and understanding of railway environment requires a large number of actual data to develop, test and verify the algorithm. Few rail transit datasets are available at present, and only RailSem19 is publicly available. Real-time semantic segmentation of railway environment using infrared thermography mainly involves the following four problems: 1) The rail transit environment is complex. 2) Infrared thermal images have more noise, low resolution, contrast and signal-to-noise ratio, and lack color features. 3) The semantic segmentation network model has a large amount of computation, which makes it difficult to meet the requirements of real-time. 4) An infrared dataset for the rail transit environment is lacking.

To solve these problems, we used the infrared thermography to collect the infrared data of Suzhou Rail Transit Line 1 in Jiangsu Province, China and proposed an infrared dataset for the rail transit – RailVID. By combining infrared thermal imaging and semantic segmentation based on deep learning, a real-time semantic segmentation method for rail transit is proposed. Keeping in mind the requirement of infrared image noise, low resolution, lack of color features, and real-time performance of the rail transit, we propose an improved BiSeNet, which is based on BiSeNet. The network is improved according to the characteristics of infrared images. The improved methods include: 1) The global maximum pooling layer is more effective in retaining the image texture information than the global average pooling layer. Therefore, the global maximum pooling layer is used to replace the global average pooling layer in the attention enhancement module and feature fusion module of the network to retain the texture details of infrared images. 2) Further fusion of the infrared image low-level features on the path of obtaining spatial information.

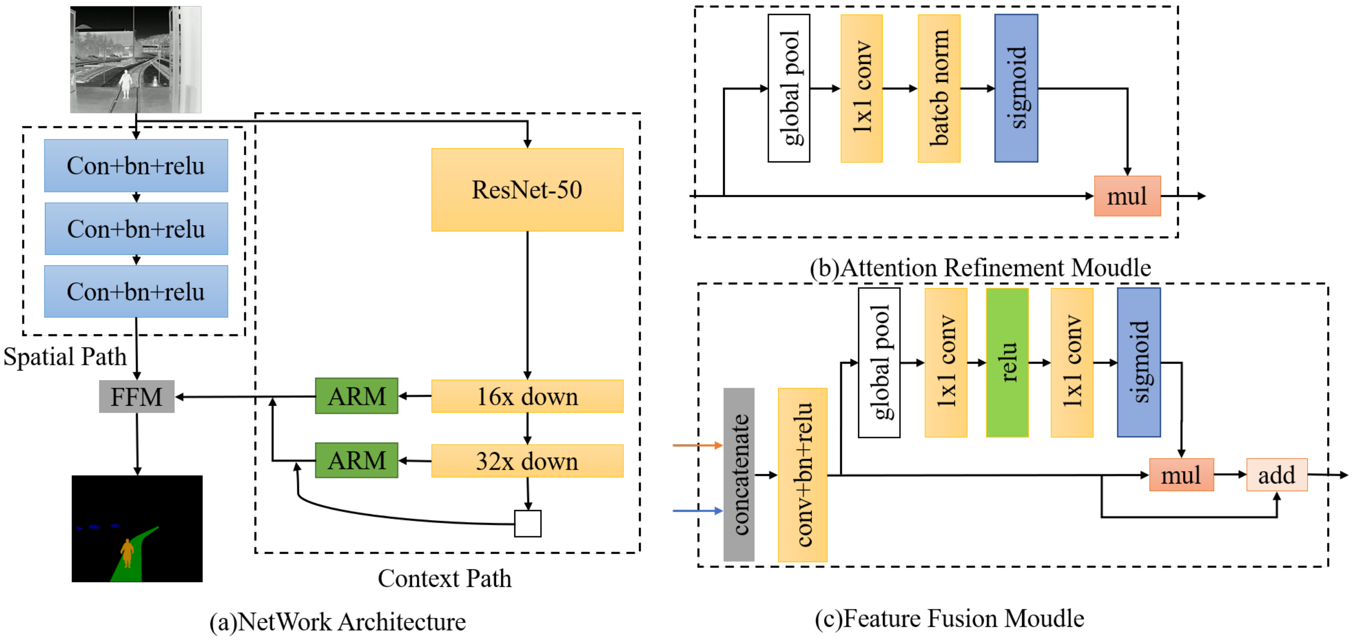


Figure 2. BiSeNet network structure.

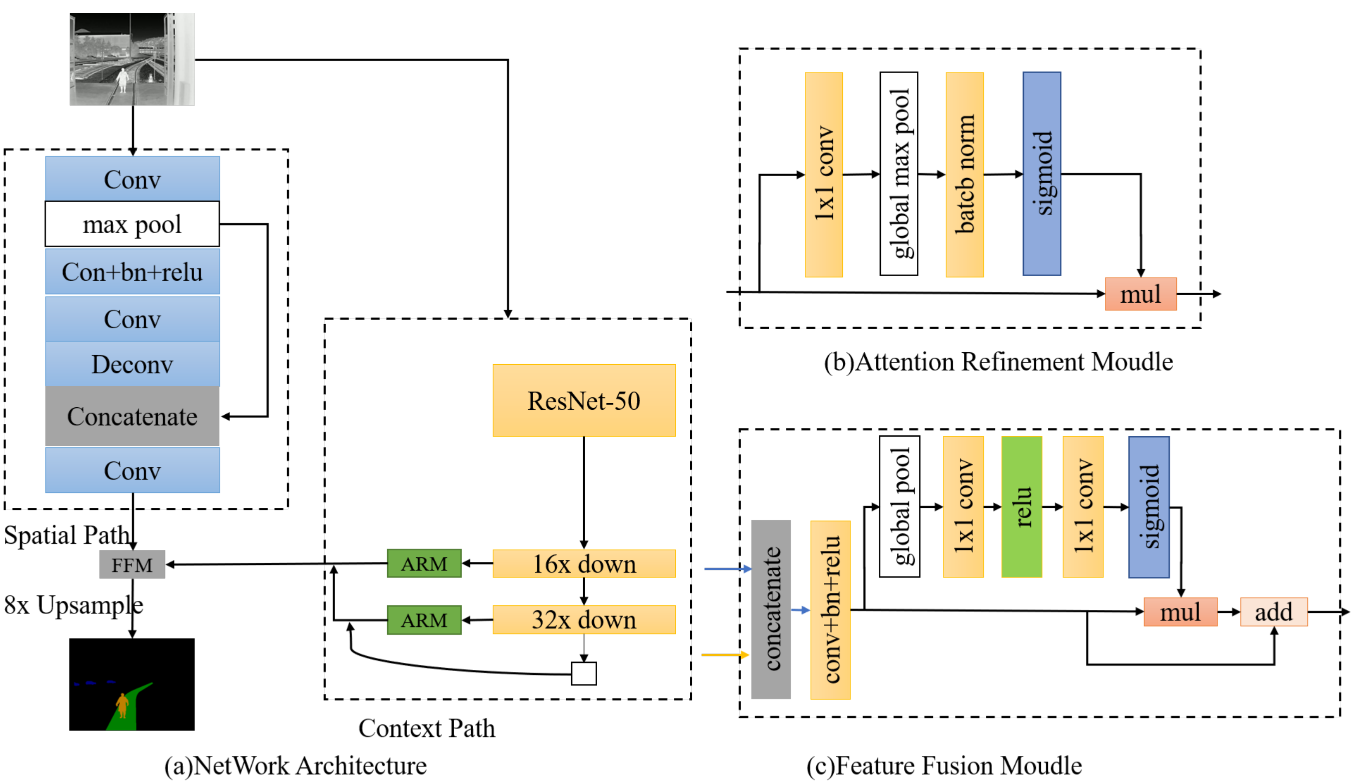
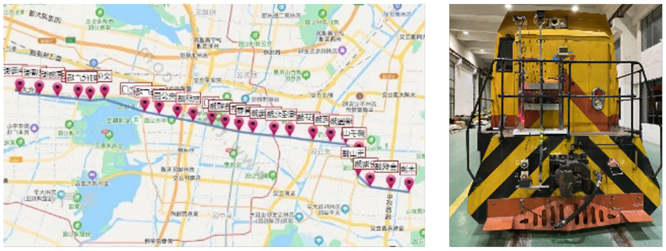


Figure 3. Improved BiSeNet network architecture.

III. Infrared Semantic Segmentation Network

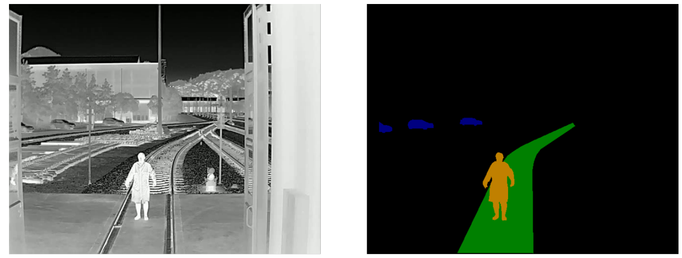
A. BiSeNet

BiSeNet network structure is shown in Figure 2. In order to achieve fast real-time segmentation without sacrificing Spatial information, the BiSeNet structure is divided into two branches: Spatial Path (SP) and Context Path (CP). The spatial path uses three convolution layers with step 2, with batch normalization and activation function Relu to output 1/8 size feature maps of the original image, and uses feature maps of



(a) Subway map (b) Subway moving platform

Figure 4. RailVID data source and acquisition platform.



Original data Label image

Figure 5. Samples of track area and surrounding obstacles.

TABLE I. CATAGORY AND CORRESPONDING RGB VALUES.

Number	Category	R	G	B
1	**background**	0	0	0
2	**railway**	0	128	0
3	**car**	0	0	128
4	**people**	64	128	0

TABLE II. NUMBER OF IMAGES IN THE TRACK AREA FOR EACH SCENE.

Route scene	Carport	Depot	Tunnel
Straight	0	55	26
Curve	16	180	22
Fork	0	165	0

TABLE III. NUMBER OF SURROUNDING OBSTACLES IN EACH SCENE.

Object scene	Carport	Tunnel
Obstacles	442	165

a larger scale to retain rich spatial information. The context path uses a lightweight model Xception to extract high-level features and adds an attention optimization module at the end of the model. The upsampling operation is omitted, and the global average pooling is directly used to capture global context information and to calculate the attention vector to obtain a maximum perception field. Finally, we use the feature fusion module to fuse the low-level information in the spatial path with the high-level information output by the context path. The specific step is to join the two output features and then normalize the scale of the balance features. The output results are transformed into feature vectors through a global pooling operation, and then a weight vector is obtained by 1×1 convolution calculation, so that the features of the two can be re-weighted.

B. Improved BiSeNet

Based on the real-time bilateral-semantic segmentation network structure, we propose an improved bilateral-semantic segmentation network structure for rail transit infrared image semantic segmentation. The structure is shown in Figure 3.

According to the idea of infrared image processing in reference [14], a method of fusing the low-level features of infrared image by adding a pool layer, deconvolution layer and full connection layer is proposed. Thus, the network can better restore the spatial resolution of infrared images and provide a larger visual field. It can enhance the attention model in the context path and the feature fusion module integrating two-way features of real-time bilateral semantic segmentation network. According to the characteristics of fuzzy details and the low contrast of infrared image, the global maximum pool layer is used to replace the global average pool layer in each pooling layer of the module network architecture to retain the texture details of the infrared image.

Considering the characteristics of infrared images and the real-time requirements of the system, the context path adopts the ResNet-50 feature extraction network, combined with global pooling. Then, it merges the intermediate results of ResNet-50(16x downsampling and 32x downsampling) and the output of global pooling.

IV. Dataset

A. Data collection

Figure 4 shows that we use the AT615X infrared thermal instrument from InfiRay to collect data, with the highest resolution up to 640×512, and 10 fps (frame per second) real-time images can be output at this resolution. Through the program we develop, the sensor data can be transmitted to the database in real-time, so as to realize data analysis and storage. We select the video frames at the proportion of 1 frame in every 10 frames to obtain all image data, and then screen the redundant images and blind areas according to the scene and train line conditions. Finally, we collect a total of 1071 images.

B. Data annotation

Due to the particularity of railway environment and the characteristics of noise and low definition of infrared imaging, we focus on the simple track area, people, and cars around the track area, as seen in Figure 5. Therefore, we mark the track area and classify people and cars using Labelme tool for manual annotation. When labeling segmentation categories, semantic segmentation tasks require that masks of different colors should be assigned to different categories to be segmented. Therefore, RGB values corresponding to Mask labels of each category with different colors are shown in Table I.

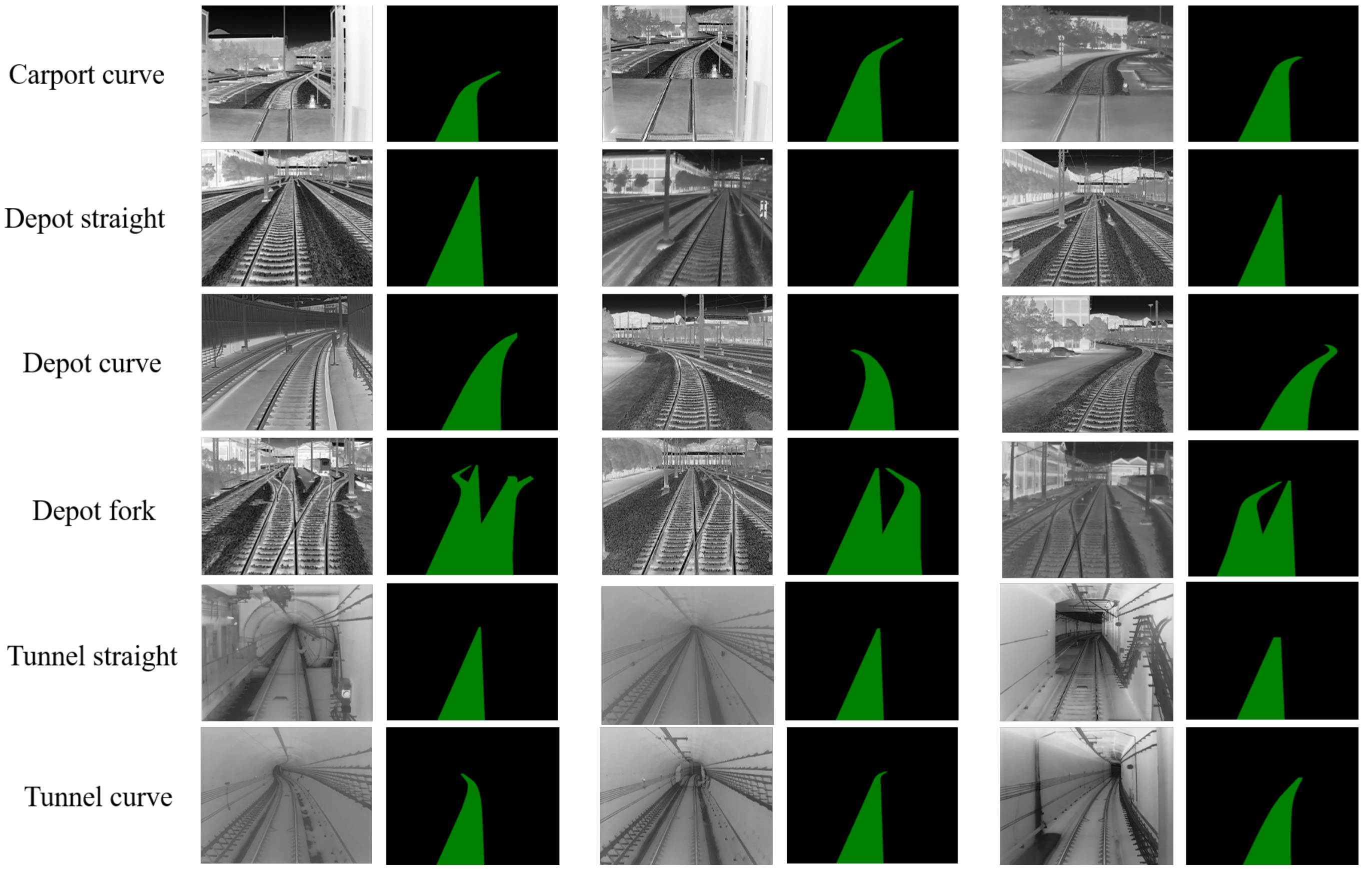


Figure 6. Track area original image and track area annotation example.

C. Data partitioning

According to the application scene, RailVID is divided into two parts: track area and obstacle, and a total of 1071 images are selected. Some examples of track area original images and track area annotation are shown in Figure 6. For the track area, we divide it into three parts: carport, depot and tunnel, and only mark the track area. The number of route scenes are shown in Table II.

There are many kinds of elements with high complexity in the railway environment. We also segmented other key elements to improve the perception ability of train perception and enrich the semantic information of railway scenes. Similarly, we divided the scene into two parts: carport and tunnel, and marked the track area, people and cars. The result are shown in Table III.

V. Experiment and analysis

A. Model training

The specific software/hardware configuration of the improved RiSeNet we proposed is shown in Table IV. The ResNet-50 pre-training model was used to initialize some parameters. The RMSprop optimizer was used to optimize all parameters. The learning rate was set to 0.0001 and the attenuation rate was set to 0.995. The parameters of each layer of the deep network were randomly initialized. The training batch was set to 100, and the optimal semantic segmentation model was generated after training.

TABLE IV. HARDWARE AND SOFTWARE CONFIGURATION.

Configuration	Hardware / software
Platform	Win10
Environment	Tensorflow1.12.0+CUDA9.0
CPU	Core i7-8750H
GPU	Geforce GTX 950

B. Evaluation metrics

To verify the reliability of the method we proposed for infrared image segmentation of track area and surrounding areas, we assumed a total of $k+1$ classes (from L_0 to L_k, including an empty class background or a class). P_{ij} is the number of pixels that belong to class I but are divided into class J, P_{ji} is the number of pixels that belong to class J but are predicted to be class I, P_{ii} for real pixel number. The indicators used to evaluate real-time semantic segmentation are as follows:

1) Mean Intersection over Union ($mIoU$), the intersection of the predicted region and the actual region divided by the union of the predicted region and the actual region.

$$mIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{P_{ii}}{\sum_{j=0}^{k}P_{ij}+\sum_{j=0}^{k}P_{ji}-P_{ii}} \quad (1)$$

 ISBN: 978-1-61208-941-6

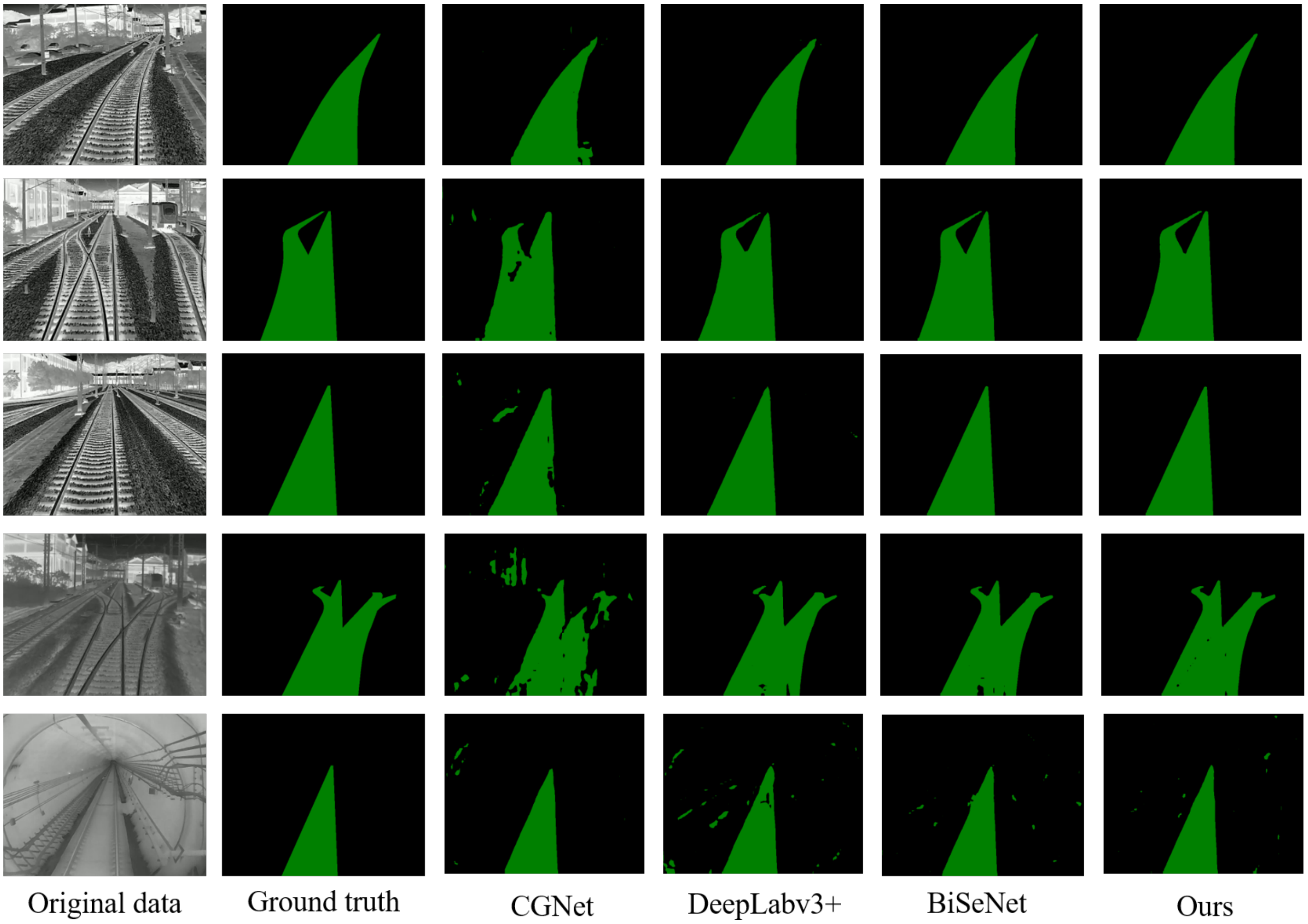


Figure 7. Comparison of effects of different semantic segmentation methods.

TABLE V. COMPARISON OF PERFORMANCE OF DIFFERENT METHODS.

Network	PA/%	mPA/%	mIoU/%	FPS/$f \cdot s^{-1}$
CGNet	60.33	51.17	59.24	53.0
DeepLabv3+	76.12	74.03	75.29	12.0
BiSeNet	78.01	76.54	77.57	45.2
ours	84.45	82.36	82.94	40.0

2) Pixel Accuracy (PA), the proportion of correctly labeled pixels to the total pixels.

$$PA = \frac{\sum_{i=0}^{k} P_{ii}}{\sum_{i=0}^{k}\sum_{j=0}^{k} P_{ij}} \tag{2}$$

3) Mean Pixel Accuracy (mPA), the average Pixel Accuracy of all classes.

$$mPA = \frac{1}{k+1}\sum_{i=0}^{k} \frac{P_{ii}}{\sum_{j=0}^{k} P_{ij}} \tag{3}$$

4) Frames Per Second (FPS), which measures the real-time performance of the semantic segmentation algorithm.

$$FPS = \frac{N_f}{T} \tag{4}$$

In formula(4), N_f is the number of video frames, and T is the consumption time, in unit s.

C. Method comparison

To verify the advantages of the proposed method in infrared image segmentation of track area, we compared with the existing three representative semantic segmentation frameworks in visible environment (CGNet, DeeplabV3+, BiSeNet) over the same dataset with the same training parameters to test the models.

In the comprehensive test, infrared images of the track area taken by real shots in the test set were used for comparative testing. The comparison is made using pixel accuracy (PA), average pixel accuracy (mPA), Mean Pixel Accuracy (MIoU) and frame per second (FPS). The comparative test results of the method we proposed compared with the other three methods are shown in Table V.

As can be seen from the comparative test index analysis in Table V, in terms of real-time performance, the improved BiSeNet is lower than BiSeNet due to the enhanced network depth. According to literature [4-5], if the FPS is above 25 frames /s, the algorithm can be identified as having real-time properties. Therefore, our method meets the real-time requirements of the semantic segmentation. The improved BiSeNet network, however, has obvious advantages over other three semantic segmentation methods in Pixel Accuracy (PA), Mean Pixel Accuracy (mPA) and Mean Intersection over Union ($mIoU$). For example, compared with BiSeNet network, our method improves the PA by 6.44 %, the mPA by 5.82 % and the MIoU by 5.37 %.

The improved BiSeNet network is optimized for retaining infrared image features and can retain more details of infrared image semantic segmentation in the track area and its surroundings. By comparing the infrared image segmentation results of the track area environment in the test set and the other three methods, it can be seen from Figure 7 that the segmentation effect of the improved BiSenet network is better and closer to the real annotation image.

VI. CONCLUSION AND FUTURE WORK

This paper described in detail about RailVID, a new infrared image dataset for rail transit that focuses on the intelligence and autonomy of rail transit train operation. We believe it will drive further development in the field of fully automated rail transit train operation. At the same time, we propose an improved BiSeNet real-time semantic segmentation network by fusing the low-level features of infrared image by adding a pool layer, a deconvolution layer and a full connection layer and replacing the global average pool layer with the global maximum pool layer for rail transit in the complex rail transit environment. On the collected dataset, the method achieves 82.94% of $mIoU$ and 40 f/s of FPS on the collected dataset, which satisfies the real-time semantic segmentation of rail transit. In the future, we will make more optimizations for the dataset and expand more data types, while a more efficient semantic segmentation network will be used to improve the characteristics of infrared images to achieve a better real-time semantic segmentation effect. This research is committed to providing the best method for fully automatic train operation in the rail transit scene and finally completing the construction of the feasibility platform.

ACKNOWLEDGEMENT

This work is supported by Scientific research project funding of Suzhou Rail Transit Group Co., Ltd. : SZZG06YJ6000017, thanks to corresponding author Cheng Wu, let me have the opportunity to participate in the meeting. Yuan and Mei contribute equally to the article.

REFERENCES

[1] B. Yum, "A Study on Railway Safety System Based on Accident Analysis," Journal of The Korean Institute of Plant Engineering, vol. 20, no. 1, pp. 101–106, 2015.
[2] J. Xiao, "Analysis of traffic accidents in Urban Rail Transportation," Information Week, vol. 1, no. 6, pp. 71–71, 2019.
[3] J. Long, E. Shelhamer, and F. Darrell, "Fully Convolutional Networks for Semantic Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640–651, 2015.
[4] Z. Zhao, W. Chen, X. Wu, P. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," LET Intelligent Transport Systems, vol. 11, no. 2, pp. 68–75, 2017.
[5] U. Safder, K. Nam, D. Kim, S. Heo, and C. Yoo, "A real time QSAR-driven toxicity evaluation and monitoring of iron containing fine particulate matters in indoor subway stations," Ecotoxicology and Environmental Safety, vol. 169(MARa), pp.361–369, 2018.
[6] V. Badrinarayanan, A. Kendall, and R. Cipolla, "A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, pp. 2481–2495, 2019.
[7] F. Yu and V. Koltun, "Multi-Scale Context Aggregation by Dilated Convolutions," Computer Science arXiv:1511.07122v1.
[8] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," Computer EnCE no. 4, pp. 357–361, 2014.
[9] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell vol. 40, no. 4, pp. 834–848, 2018.
[10] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," arXiv preprint arXiv:1706.05587, 2019.
[11] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," In: Proceedings of the 15th European Conference on Computer Vision, pp. 833–851. Springer International Publishing, Switzerland, 2018.
[12] T. Wu and S. Tang, "CGNet: A Light-weight context guided network for semantic segmentation," IEEE Conference on Computer Vision and Pattern Recognition. Intell., vol. 30, pp. 1169–1179, 2019.
[13] R. Barth, J. Ijsselmuiden, J. Hemming, E. Henten, and J. Van, "Synthetic bootstrapping of convolutional neural networks for semantic plant part segmentation," Computersand Electronics in Agriculture, Intell vol. 161, pp. 291–304, 2019.
[14] Z. He, Y. P. Cao, Y. F. Dong, J. X. Yang, and Y. L. Cao, "Single-image-based nonuniformity correction of uncooled long-wave infrared detectors: A deep-learning approach," Applied Optics. Intell., vol. 57, no. 18, pp. 155–164, 2018.
[15] Y. H. Chen, N. Zhu, Q. Wu, C. Wu, W. L. Niu, and Y. M. Wang, "MRSI: a multimodal proximity remote sensing data set for environment perception in rail transit," Int J Intell Syst, pp. 1–27, 2021.
[16] Z. Chen, C. Wu, L. J. Zhang, and Y. M. Wang, "Near Real-time Situation Awareness and Anomaly Detection for Complex Railway Environment," IEEE Conference on Cognitive and Computational Aspects of Situation Management, pp. 1–8, 2021.
[17] G. Lin, A. Milan, C. Shen, and I. Reid, "RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5168–5177, 2017.

Systems Governance in Joint Military Acquisition Requirements Generation: A Systems Thinking Approach

Jack Goldberg
Stevens Institute of Technology
Hoboken, NJ 07030 USA
email: jgoldbe3@stevens.edu

Mo Mansouri
Stevens Institute of Technology
Hoboken, NJ 07030 USA
email: mmansour@stevens.edu

***Abstract*—Acquisition latency has been a topic of interest for both systems researchers and the United States government. Using a systems thinking approach, we argue that current reform efforts for the Joint Capabilities Integration and Development System (JCIDS) wrongly focus on process improvement when latencies are more closely tied to inadequate systems governance.**

Keywords-systems thinking; systems dynamics; acquisition; governance.

I. INTRODUCTION

House Armed Services Committee report 116-120 ordered an assessment of the timeliness of the Joint Capabilities Integration and Development System (JCIDS) in creating and validating requirements documents for the US military [1]. JCIDS, the process by which military capability gaps are addressed and turned into requirements documents, has been marred by inefficiencies, thus prompting the assessment. It takes far too long for warfighter effectiveness. As a result, military services have opted to pursue alternate capability development pathways, notably the Middle Tier of Acquisition (MTA), as a way to rapidly prototype solutions and circumvent JCIDS entirely. This is problematic, as JCIDS is meant to ensure interoperability between service branches. Joint operation is an increasingly important function of the US military and breakdowns can prove catastrophic.

In summary, JCIDS consists of a series of document writing and review activities that result in a validated document that can be used to request proposals from private military contractors. Military services submit their capability requirement documents to JCIDS which, along with representatives from other services, review it to make sure the capability is interoperable across the joint warfighter. The submitting service may adjudicate comments offered by the reviewers. After a series of reviews and rewrites, a validated requirement document emerges that describes a military system that will maintain interoperability.

The system was developed in response to the interoperability failures in the joint military effort in the early 2000s. For example, the case of the Second Battle of Fallujah, or Operation Al Fajr (November - December 2004), highlights the danger of interoperability breakdown. During this episode, Army and Marine forces were working together to destroy enemy targets in the city. However, the two forces struggled to communicate with each other due to incompatible communication technologies: the Army used radios while the Marines primarily used Internet chat like Microsoft Chat [2].

The governance issues in JCIDS perpetuate a distrust of the JCIDS process. Service branches refuse to engage with JCIDS because they do not believe that their capability gaps will be addressed fast enough.

Using a systems thinking approach, we fully characterize JCIDS and its issues related to process latency. By considering the whole of JCIDS and its relation to other systems, instead of breaking it down into its constituent parts, we capture a richer depiction of the system. We argue that the latencies in JCIDS are not a result of a poorly designed process, but rather come from less obvious forces like misaligned incentives or poor organizational culture. Thus, we believe that the House Armed Services Committee is misled in requesting a review of the process and not the actual sources of latency.

The rest of the paper is structured as follows. In Section II, we present a literature review that addresses the use of systems thinking as well as other research on military acquisition. In Section III, we describe the methodology used for this work. In Section IV, we describe and map the relevant stakeholders of JCIDS. Section V makes account of the value adding processes in JCIDS. Next, Section VI analyzes the shaping forces of JCIDS. Section VII presents two conceptual models, a systemigram and systems dynamics representation of JCIDS. In Section VIII, we synthesize the information into an analysis of the system as a whole. Then, in Section IX, we contextualize the analysis within the topic of systems governance and JCIDS latency. Finally, in Section X, we offer concluding remarks and suggest future research.

II. LITERATURE REVIEW

Systems thinking has been useful in assessing many complex systems. Boardman et al. use a systems thinking approach to characterize enterprise resilience in maritime system of systems [3]. The approach is also amenable to different scenarios and at many different levels of abstraction. Along with maritime systems, systems thinking has been used to develop a framework for energy behavior in smart cities [4].

Relevant to this work, systems thinking has also been used to analyze military acquisition. Assidmi et al. use a systems thinking approach to incorporate human-centered factors into projections of cost growth in weapon system acquisitions [5].

 ISBN: 978-1-61208-941-6

Finally, the concept of governance is very relevant to systems thinking as emergent behavior is influenced by governance structures. Systems thinking is used to identify four domains that impact public trust in healthcare: medical errors and malpractice lawsuits, the roles of third-party beneficiaries in medical lawsuits and the conflicts presented on mass and social media, the public trust of healthcare services, and the healthcare quality improvement efforts [6]. These public health outcomes are dependent on systems governance. Research on systems governance is still in its infancy and lacks a robust literature.

We seek to enrich the literature by providing a systems thinking analysis of governance within a complex enterprise system. We use JCIDS as a case study in our work.

III. METHODOLOGY

We use a combination of document analysis and interviews to pursue our systems thinking approach. Much of the details on the operation, including the required documents and reviews, of JCIDS were taken from the JCIDS manual, Manual for the Operation of the Joint Capabilities Integration and Development System [7].

Next, we conducted a series of unstructured interviews with acquisition experts. The interviews were coded for common themes. The insights from the interviews were critical in the development of our systems thinking analysis. The interviews help in developing an understanding of JCIDS that is not addressed in official documents, like sources of latency or collaborative culture.

IV. STAKEHOLDERS

First, we identify the stakeholders in JCIDS. These actors are the most important in shaping the behavior of the system. Fig. 1 lists these stakeholders, their relationship to JCIDS, and a description.

TABLE 1. NOTABLE STAKEHOLDERS IN JCIDS.

Name	Type	Description
Gatekeeper	Active	The key facilitator of JCIDS and the first to receive documents from the sponsor.
FCB	Active	The Functional Capabilities Board is a review body that is dedicated to a particular area of military capability.
JROC	Active	The Joint Requirements Oversight Committee is the ultimate approval body in JCIDS.
Sponsors	Active	The sponsor of an acquisition project drafts the JCIDS documents.

A key consideration is how the stakeholders interact with each other. The behaviors of the stakeholders are shaped by their personal incentives. The incentives can cause reinforcing behavior or mitigating behavior. These interactions define the emergent behaviors of a system and ultimately shape the success of the system. In terms of JCIDS, a hierarchical military system, there are prescribed behaviors. Fig. 2 illustrates the stakeholder interactions that occur in JCIDS.

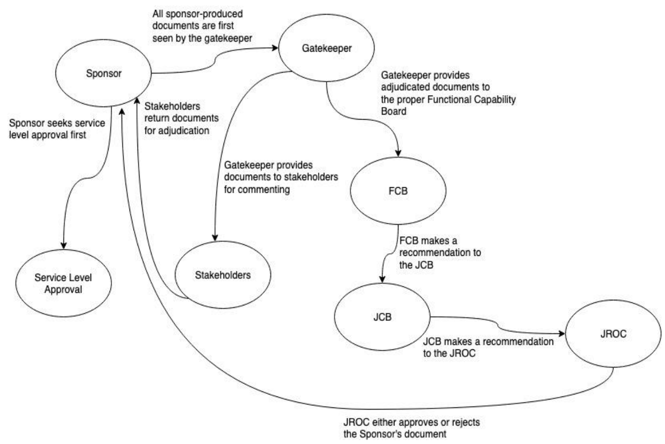


Figure 1. Stakeholder interaction map within JCIDS.

V. VALUE ADDING PROCESSES

Next, we address the idea of added value in JCIDS. First, we will consider the factors that add value to JCIDS. These things help JCIDS accomplish its goal of efficiently validating requirements documents.

A. *Information Technology*

Information technology facilitates the requirement validation process. JCIDS staff use Knowledge Management/Decision Support (KM/DS) to coordinate and archive requirements documents. However, the KM/DS system often has missing and incorrect data points.

B. *Training*

Training is key in empowering JCIDS personnel. Knowing how to write and comment on requirements documents results in quicker validation times because less adjudication will be needed.

C. *Experience*

Similar to training, experience empowers JCIDS personnel. Experience within JCIDS makes JCIDS staff more knowledgeable about document writing and also reduces confusion. Experienced personnel knows exactly how JCIDS is supposed to operate. Yet, there is rapid personnel turnover in JCIDS.

D. *Warfighter Goal Clarity*

When warfighter goals are clearly defined, JCIDS approval bodies can easily determine whether an acquisition project is necessary. In this situation, acquisition projects will not even be pursued if they do not align with warfighter needs.

The purpose of JCIDS is to add value to the military acquisition system. After considering the variables that add value to JCIDS, we consider the things that JCIDS adds value to.

E. *Warfighter*

JCIDS is supposed to address capability gaps and enhance joint capability. This will result in better outcomes on the battlefield

F. *"Little a" Acquisition*

Considering the output of JCIDS is the input of contracting, better defined projects help with project contracting. Well defined requirements are useful for contractors that are making bids.

VI. SHAPING FORCES

JCIDS operates in a well-defined environment. It is a part of the US military acquisition ecosystem, which consists of itself, budgeting, and "little a" acquisition. Because it is so well bound, it is relatively easy to identify the shaping forces of JCIDS. We start by identifying the internal forces that are not explicitly mentioned in the JCIDS manual, but still impact acquisition outcomes.

A. *Culture*

Organizational culture is important in determining acquisition outcomes. One facet of culture is risk tolerance. JCIDS, a risk averse organization, will take a considerable amount of time in making sure requirements documents are very well defined. Further, a culture of complacency in JCIDS does not inspire innovation. Personnel are satisfied with the latencies within the system.

B. *Collaboration*

Collaborative design is commonplace in the private sector. It shapes project outcomes by incorporating multiple perspectives throughout the design process. There is not a uniform practice of collaboration in JCIDS. The Army utilizes Cross-Functional Teams [8] while the other services are less collaborative. These differences in collaborative structure impact acquisition timelines.

After analyzing the internal shaping forces of JCIDS, we move onto the forces that affect the environment that JCIDS operates in, and induce action. These forces influence either the inputs to JCIDS or help dictate what an output looks like.

C. *Adversarial Developments*

Adversarial developments refer to the changes in threats to US security. This can be the emergence of a new threat, like an antagonistic state, or a new capability among existing US enemies. These developments inspire a reaction by the US military, which ultimately results in a JCIDS process.

D. *Innovation*

Private sector or academic innovation pushes the state of the art, and results in a military capability gap. Many US adversaries can adopt innovation very quickly as well, further hindering the US's warfighter capability. Thus, innovation spurs acquisition and, thus, JCIDS.

E. *Political influence*

Acquisition can have political value for a great number of reasons. First, politicians have an interest in national security. They should want to protect the safety of Americans. Acquisition can be useful in accomplishing that. Second, there are economic incentives for a lot of politicians. Acquisition can employ a lot of people in a politician's district. A politician may also be interested in acquisition oversight, which will force accountability in JCIDS. Political pressure can speed up the JCIDS process but is only used in certain, advantageous, cases.

F. *Budgeting*

Budgeting is a major shaping force in JCIDS because acquisition projects cannot be realized if there are no funds available to finance them. In fact, there is an affordability analysis built into the JCIDS process.

VII. CONCEPTUAL MODELS

We develop conceptual models to further our understanding of the system and how system governance impacts JCIDS effectiveness.

The following systemigram, displayed in Fig. 2., creates a series of narratives concerning JCIDS. The systemigram was created to help visualize the nature of a system [9]. The main line of the diagram shows the standard story whereby capability gaps are developed into validated requirements documents. Yet, there are also peripheral narratives regarding innovation, personnel, and leadership that are key in JCIDS.

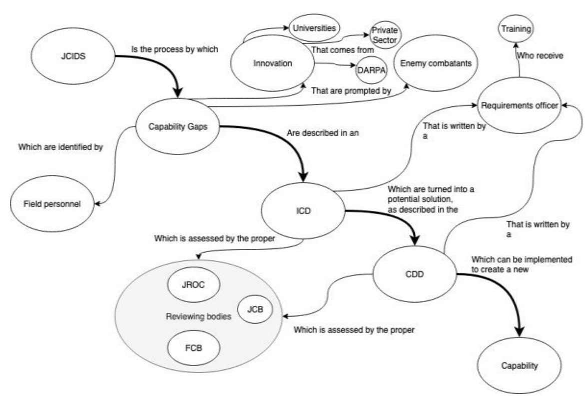


Figure 2. JCIDS systemigram.

Finally, Fig. 3 illustrates the system dynamics of JCIDS. It shows how different aspects of JCIDS reinforce and balance each other. The United States and its adversaries are in an effective arms race for better capabilities, which is only balanced by budgeting restraints. Therefore, acquisition speed is a key military priority, which is hindered by the latency of JCIDS.

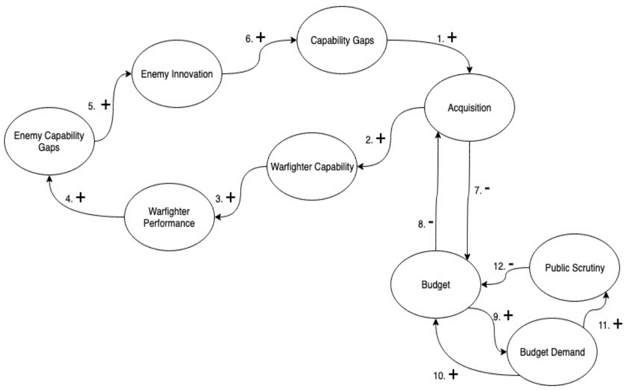


Figure 3. JCIDS systems dynamics.

VIII. SYSTEMS ANALYSIS

Our systems thinking analysis illustrates why JCIDS is ineffective in validating requirements documents. At a high level, warfighter needs require that JCIDS operate quickly; troops on the battlefield need acquisition to quickly fulfill capability gaps. The US military needs to adopt innovation at the same rate as their adversaries in order to maintain a tactical advantage. Yet, acquisition processes are the subject to oversight. The American public and Congress demand fiscal responsibility. This results in regulation and policy that slows down JCIDS. Requirements documents require a lot of reviews to ensure that the acquisition is necessary, so funds are not being misused. The pressures to be fast and deliberate at the same time are at odds with each other. This reality puts JCIDS in a precarious position. It cannot adequately satisfy both stakeholders, and often fails both.

In a system with opposing forces, an equilibrium is often reached. JCIDS operates in a unique environment because there are limited resources in the form of budgeting and the government is a monopsonist in purchasing military equipment. So, the oversight bodies, including Congress, exert the most influence on JCIDS. Therefore, the equilibrium favors the stakeholders who want a deliberate process. Yet, it is the service branches that are actually interacting with the system by creating requirements documents.

At a lower level, JCIDS does not have governance structures in place to facilitate rapid development of requirements documents. There is very little collaboration in the development of these documents, which could speed up the process. There is also an inadequate infrastructure for information technology; data collection is insufficient and often inaccurate [10].

Considering the contradictory influences in JCIDS, service branches often circumvent the process by using the Middle Tier of Acquisition. This acquisition track gives sponsors the ability to rapidly prototype and field capability without dealing with JCIDS. The faster acquisition time and reduced oversight has made MTA an attractive option. The threat of interoperability failure remains in MTA, but the consequences are delayed, which makes sponsors discount them.

IX. SYSTEMS GOVERNANCE

A crucial result of the inadequate governance of JCIDS is a lack of trust in the system. The opposing forces of speed and oversight in JCIDS indicate why there is a lack of trust among stakeholders. If JCIDS cannot field materiel fast and cheap enough, then sponsors will place their trust in other acquisition processes, like MTA, that will deliver capabilities much faster, even at the expense of interoperability.

An important consideration in addressing stakeholder trust in JCIDS is the incentives of the constituent stakeholders. Each part of the system has their own incentive and will try to maximize their own interest. Considering JCIDS has stakeholders from all service branches, joint staff, legislative bodies, and more, there are many agents pulling the system in their direction. With increased competition for resources and influence, stakeholders may develop a distrust of the system.

Furthermore, personnel development shortcomings, which are caused by lack of training and scheduled personnel rotations, hinder the development of trust among stakeholders. Since JCIDS staff and service level requirements officers are frequently leaving their departments, there is no opportunity for services to develop a trusting relationship with JCIDS. This problem is magnified when leadership turns over, as their policies are abandoned upon leaving.

X. CONCLUSION

We lay out a comprehensive analysis of JCIDS. The system is best understood by its relationships to the whole rather than taking apart its components. Through the systems thinking approach, we describe the critical stakeholders, value adding processes, shaping forces, and conceptual modeling of JCIDS.

The purpose of this work is not to suggest solutions for improving JCIDS. Rather, we seek to holistically characterize the problems in JCIDS. Often the true root of a problem is not the most visible. Our systems thinking analysis highlights some of the factors that are hindering JCIDS that may normally be overlooked, especially by those in power who have the ability to enact change. Our analysis shows that the latency in JCIDS is not wholly a result of poorly designed processes, but the culmination of poor systems governance including the misalignment of incentives, lack of technological infrastructure, and personnel shortcomings.

Future research should analyze possible solutions to overcome JCIDS inefficiency. Scenario analysis and surveys can help ground this conceptual analysis.

While it is not a well-known system, JCIDS is incredibly important in American society. It determines how American troops are equipped and impacts the public's tax burden. Understanding and improving this system is crucial to American interests both at home and abroad.

REFERENCES

[1] House of Representatives Rept. No. 116-120, at 176-177 June 19, 2019, Available: https://www.congress.gov/congressional-report/116th-congress/house-report/120/1 [last access April 2022].

[2] M. Matthews, “Operation AL FAJR: A Study in Army and Marine Corps Joint Operations”, Army Command and general Staff Coll Fort Leavenworth KS Combat Studies Inst, 2006, Available: https://apps.dtic.mil/sti/citations/ADA454930 [last access April 2022].

[3] M. Mansouri, B. Sauser, and J. Boardman. “Applications of systems thinking for resilience study in maritime transportation system of systems”, The 3rd Annual IEEE Systems Conference, pp. 211-217, 2009.

[4] N. Khansari, H. R. Darabi, M. Mansouri, and A. Mostashari. “Case study of energy behavior: Systems thinking approach”, The Annual IEEE Systems Conference (SysCon) Proceedings, pp. 373-376, 2015.

[5] L. Assidmi, S. Sarkani, and T. Mazzuchi. “A systems thinking approach to cost growth in DoD weapon systems”, IEEE Systems Journal, vol. 6 no. 3, pp. 436-443, 2011.

[6] P. Kaewkamjonchai et al., “A systems thinking approach to understanding public trust in healthcare services and doctor-patient relationship in the contexts of medical errors”, ResearchSquare. Preprint. Oct 05, 2020. [last access April 2022] 10.21203/rs.3.rs-77431/v1

[7] Department of Defense, “Manual for the Operation of the Joint Capabilities Integration and Development System”, 2018, Available: https://www.acq.osd.mil/asda/jrac/docs/2018-JCIDS.pdf [last access April 2022]

[8] D. A. Reisinger. "Cross-Functional Teams for the Future." *Army Law*. 2020: 92.

[9] B. Clegg and J. Boardman, "Process integration and improvement using systemic diagrams and a human-centered approach.", *Concurrent Engineering* 4.2: pp. 119-136, 1996.

[10] Government Accountability Office. “Joint Staff Lacks Reliable Data on the Effectiveness of Its Revised Joint Approval Process”, 2021, Available: https://www.gao.gov/assets/gao-22-104432.pdf, [last access April 2022]

Systemigrams for PESTEL Analysis of an Offshore Windfarm System

Yayun Chen
Department of Science and Industrial Systems
University of South-Eastern Norway
Kongsberg, Norway
e-mail: edithchen03@gmail.com

Mo Mansouri
Department of Science and Industrial Systems
University of South-Eastern Norway
Kongsberg, Norway
e-mail: mo.mansouri@usn.no

***Abstract*—As energy transition is ongoing, many companies are looking for new opportunities in new markets. Offshore windfarm systems, especially those with large-scale hydrogen storage functions, have become an important part of the global energy transition. Companies often use PESTEL (Political, Economic, Social, Technological, Environmental and Legal) Analysis to get a macro picture of an industry environment when deciding whether to join an opportunity. However, the PESTEL Analysis analyzes each aspect one by one, often missing their interdependencies. This makes it more difficult to discover all relevant factors; moreover, the results of the PESTEL perspectives may not be well integrated. In this paper, the authors use Systemigrams as a tool to integrate the individual perspective from PESTEL aspects into a holistic understanding of an offshore windfarm system for Faroe Islands. Using this tool, management can gain a comprehensive understanding of a renewable energy system for the Faroe Islands that will inform their decision on whether to bid for the project.**

Keywords-system thinking; PESTEL Analysis; Systemigrams; offshore windfarm; Faroe Islands.

I. INTRODUCTION

This section includes three parts: first, the background of the case project; second, explaining what the PESTEL Analysis is; thirdly, the case system introduction.

A. Background

The Faroe Islands is one of the world's leading producers of renewable electrical power, with 50% of its electricity is derived from renewable energy sources. The other 50% comes from fossil fuels. Denmark aims to become 100% independent of fossil fuels by 2030 [1]. Fig. 1 shows that the Faroe Islands are located in the North East Atlantic Ocean, where the average wind speed is above 11 m/s from October to April [2]. This is ideal for producing clean renewable wind power. However, in June and July, the wind speed drops below 8 m/s, which means that a hybrid system is needed to balance the low wind season.

An offshore windfarm system is being developed to help the Danish government to achieve the goal by 2030. A project team has been developing the concept, and the next phase is convincing the company`s management that it is the ideal solution and using the concept to bid the project. During the project, the team focuses on a holistic approach that includes the Political, Economic, Social, Technological, Environmental and Legal aspects of the system of interest.

Figure 1. Location of Faroe Islands [3]

B. PESTEL Analysis

A PESTEL Analysis is an analysis of the external macro environment (big picture) in which a business operates. These are often factors which are beyond the control or influences of a business, however, are important to be aware of when doing product development, business or strategy planning [4]. It is important to understand that all external factors of the environment need to be checked and identified before making a decision. PESTEL Analysis allows to determine and analyze the key factors of change in business environment [5]. The PESTEL aspects are shown in Fig. 2.

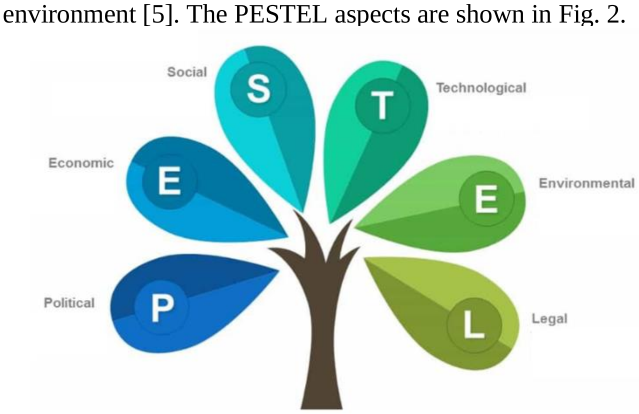


Figure 2. PESTEL aspects [6]

- Political aspects refer to the influences from the government. They include all the policies, laws and restrictions.
- Economical aspects have a major influence on how organizations run business and how profitable they are [5]. they include the taxes, duties, exchange rates, cost of living and Gross Domestic Product (GDP), etc.
- Social aspects refer to the size and growth rate of the population, immigration rates, lifestyles, education levels, and attitudes regarding investing, etc.
- Technological factors have become a huge influence for organizations in assessing and listing issues that could have a potential impact on its operations and that could be critical to its long-term future [5].
- Environmental factors relate to the influence of the surrounding environment and the impact of ecological aspects [5]. They include the weather, environmental policies, climate and support for renewable energy, etc.
- Legal aspects include, for example, employment laws, health and safety laws, taxation, etc.

It is very important to understand the offshore windfarm project from the PESTEL aspects` perspective. However, the results of the PESTEL perspectives may not be well integrated. In this context, it is necessary to apply a system thinking approach to integrate these perspectives into a comprehensive view of the offshore windfarm system.

C. Offshore windfarm system for Faroe Islands

The offshore windfarm system is a sustainable renewable energy system that uses offshore wind energy to generate electricity and extract green hydrogen from seawater. The system will supply electricity to the Faroe Islands without interruption year-round, making the Faroe Islands 100% fossil fuel free by 2030.

The system includes a wind turbine system, an electrolysis system, a hydrogen storage system, a fuel cell system, and a seawater treatment system. When the wind is active, the power from the wind turbines is fed directly to the grid. During periods of excess electricity available, electricity will be used to split water into hydrogen, which will be stored in storage tanks onshore, with a portion of the hydrogen distributed to transport fuel and another portion of the hydrogen converted into electricity again when there is no or low wind. An overview of the system is shown in Fig. 3 [7].

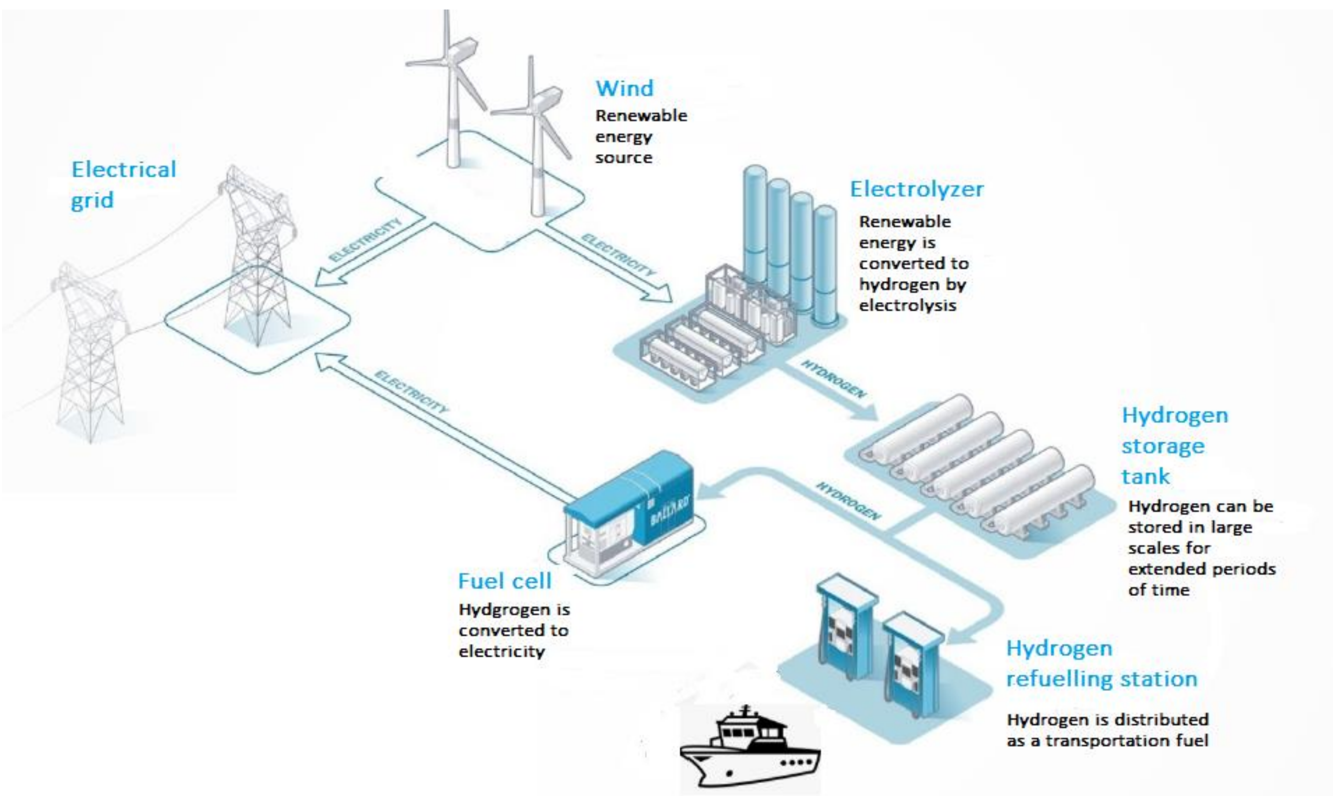


Figure 3. Offshore windfarm system overview

 ISBN: 978-1-61208-941-6

II. SYSTEM THINKING APPROACH

In this section, the purpose of using Systemigrams is explained. Then, we describe how Systemigrams are used to integrate the PESTEL aspects of the Faroe Islands offshore windfarm system, and thereby address project challenges and opportunities.

A. *Systemigrams*

Systemigrams provide a powerful tool for the analysis of systems first described in written form. Using Systemigrams to integrate the PESTEL, especially considering the interdependencies between sub-aspects, is very useful for a project to establish a holistic view of the system. Because PESTEL Analysis analyzes all aspects one by one, a Systemigram can provide a verification platform that helps analysts ensure that all relevant PESTEL aspects have been discovered. All individual aspects can be integrated through a Systemigram. This can provide a common basis for group discussions [8] and presentation to the management.

Fig. 4 shows the Systemigram of the offshore windfarm system in terms of PESTEL. From top left to bottom right is the mainstay, which is the purpose of the system. This renewable energy system will be able to produce electricity and hydrogen for consumers year-round. This will make the Faroe Islands 100% independent of fossil fuel. Furthermore, the PESTEL aspects are developed around the mainstay and are categorized by different colors.

- Blue represents political aspects. The political situation in Denmark and the Paris Agreement will both be considered. For example, the Danish government`s goal of having green energy supply all electricity in the Faroe Islands by 2030 shows the Danish government`s support for the Paris Agreement. In addition, tax laws such as the CO2 tax are DKK (Danish Krones) 173 per ton of CO2 [9]. From this, we can conclude that this system will help to save huge CO2 taxes compared to fossil fuel systems.
- Economic aspects are shown in light green. The most important economic parameters of the system are CAPEX (Capital Expenditure), OPEX (Operating Expenses), and Consumables [6] that make up the energy cost. The cost is high, however,

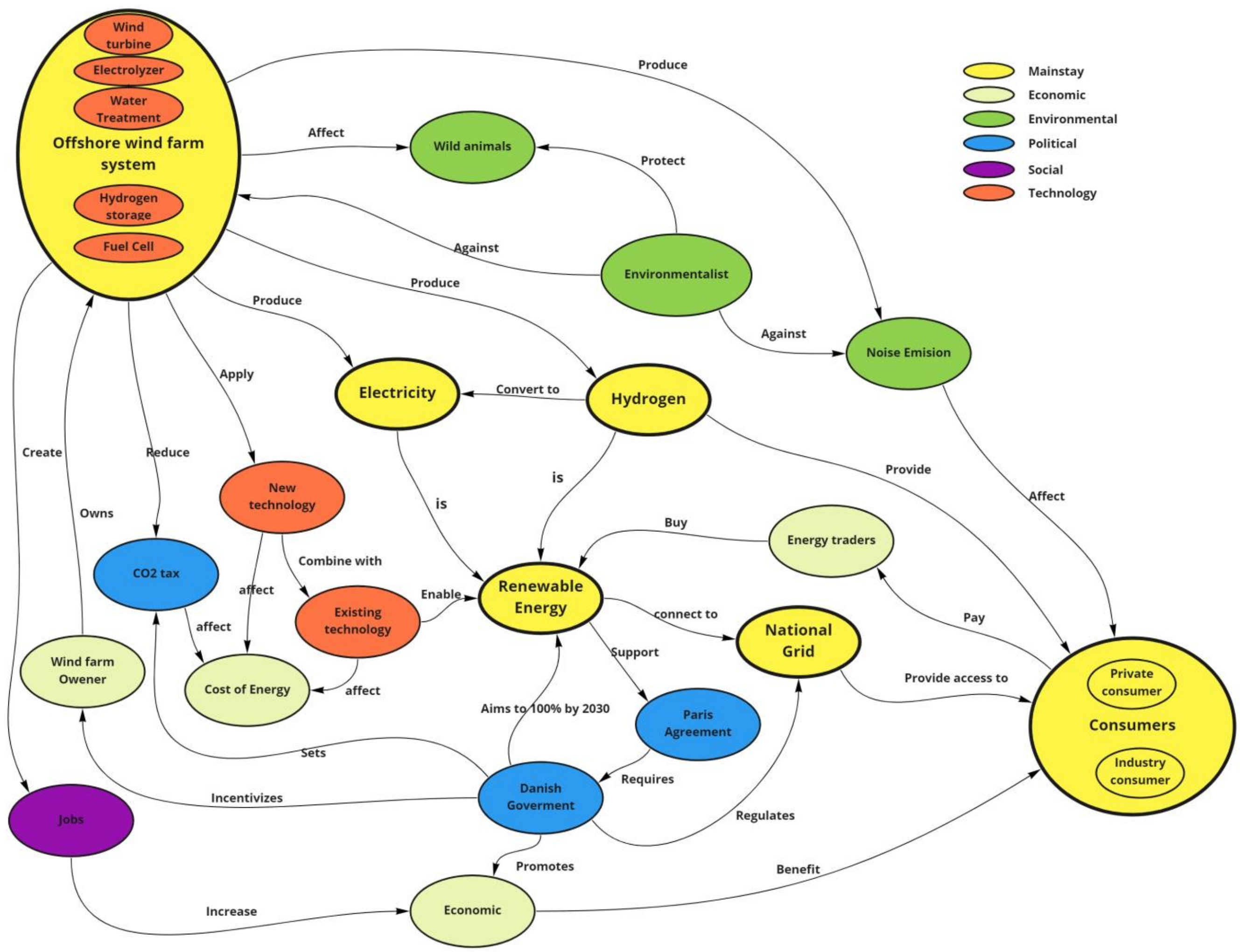


Figure 4. Systemigram of offshore windfarm

the Danish government will incentivize the renewable energy project, which will lower the project`s costs. At the same time, the CO2 tax impacts the cost significantly.

- The system will combine the existing technology with the new technology. Offshore windfarm is expected to be existing technology, as several large-scale offshore wind farms have already been deployed in the UK. However, due to the explosive nature of hydrogen, hydrogen storage presents significant risks. This requires new technology to improve it. Furthermore, since the location is new, some new knowledge input may be required during installation. These will affect the cost of energy. Despite that, using new technologies will increase the efficiency of the system, thereby increasing the output of energy. On the other hand, new technologies will make the system safer and have less impact on the environment.
- The project will create job opportunities for the local community, thereby increasing the local economy. This will benefit the residents who are consumers of the system.
- Environmentalists are normally against the windfarm project because wind turbines take up the space where the wildlife lives. However, there is clear evidence that this energy source is much safer than other methods. In addition, weather conditions will be considered as an important aspect of the environment. As mentioned earlier, wind speeds in the Faroe Islands are ideal for wind turbines all year round except in summer.
- All operations and activities follow Danish regulations.

B. Challenges and opportunities

The current challenge for the project is that the initial cost of the system is higher than traditional energy sources, and new technologies can lead to overruns. Meanwhile, there are still some impacts on local wildlife. But it is also full of opportunities, for example, this project will provide the Faroe Islands with a stable supply of renewable energy. The hydrogen will be able to balance electricity consumption throughout the year, in addition to fueling ships. Because of this project, Faroe Islands to be seen as a pioneer in clean energy. The company also will become a pioneer of the offshore windfarm system with large scale hydrogen storage and has a great opportunity to get more projects.

III. Conclusion

In this paper, a system thinking approach as a Systemigram has been applied to an offshore windfarm project to integrate individual PESTEL aspects into a holistic view of the system. At the same time, the opportunities and challenges have also been addressed. These will make the presentation to management more convincing.

Further application of Systemigrams in every aspect will be more in-depth. A Causal Loop Diagram will be created for the system of interest and the balance of dynamic forces will be examined.

References

[1] Danish Ministry of Energy, Utilities and Climate, “Denmark’s National Statement to the COP23” . [Online]. Available from: https://unfccc.int/process-and-meetings/conferences/un-climate-change-conference-november-2017/events-and-schedules/high-level-segment/high-level-segment-statements-of-cop23/cmp13/cma12. 2022.02.01

[2] D. Katsaprakakis, E. Dakanali, and K. Tzirakis, “Faroe Islands: Towards 100% R.E.S penetration” Conference on Clean Energy in European Islands, April. 2018, pp. 6.

[3] D. Katsaprakakis, E. Dakanali, and K. Tzirakis, “Faroe Islands: Towards 100% R.E.S penetration” Conference on Clean Energy in European Islands, April. 2018, pp. 2.

[4] M. Morrison, “Introduction to the PESTEL Analysis tool”. [Online]. Available from: https://rapidbi.com/Introduction-to-the-PESTLE-analysis-tool/ 2022.02.01

[5] E. Shatskaya, M. Samarina, and K. Nekhorosheva, “PESTEL Analysis as a tool of strategic analysis in international markets”, 2nd International Conference on Science and practice: a new level of integration in the modern world, Dec. 2016, pp. 48-51.

[6] E. Syverud, “New Energy Systems” University of South-Eastern Norway (USN). 2021.

[7] R. Segev, “ Optimizing Renewable Energy Storage With Hydrogen Fuel Cells” [Online]. Available from https://blog.ballard.com/renewable-energy-storage 2020.12.17

[8] R. Edson, “ System Thinking. Applied” Applied Systems Thinking Institute (ASysT), Analytic Services INC. 2008, pp. 21.

[9] Organisation for Economic Cooperation and Development (OECD). Taxing Energy Use 2019: Country Note-Denmark.

Solving Challenges in Mental Healthcare Considering Human Factors

Nynke Meijer
Behavioural Management and Social sciences
University of Twente
Enschede, The Netherlands
n.j.meijer@student.utwente.nl

Mo Mansouri
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, USA
mo.mansouri@stevens.edu

Onur Asan
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, USA
oasan@stevens.edu

***Abstract*—The mental healthcare sector is coping with several complexities, like long access times and low trust of clients. In this paper, we investigate the causes of and possible solutions for these complexities, while considering human factors. The six root causes are overtreatment, low situational awareness of planning staff, cherry-picking, unusable procedures, limited mental care knowledge among General Practitioners (GPs), and heterogeneity of clients. To solve these, establishing a collaboration between different care providers is important. Additionally, a new digital system could improve the situational awareness of therapists and planning staff. In developing and implementing these measures, it is important to take into account the stakeholders and to evaluate the ideas regularly to ensure acceptance, usability, usefulness, and trust.**

Keywords—mental healthcare, human factors, access times, waiting lists, overtreatment, collaboration, situational awareness, trust.

I. Introduction

Access times in mental healthcare are a persistent problem, while waiting for treatment can worsen the client's condition, especially in this sector. In the Netherlands, the maximum acceptable access times are bounded by the so-called 'Treek-norms'. In the mental healthcare sector, care providers and health insurers agreed on a Treeknorm of four weeks between the application and the initial consult and an additional ten weeks between this consult and the start of the treatment. However, these norms are often exceeded [1], reducing the satisfaction and trust of the clients. This is a complex problem to solve because many stakeholders are involved and many factors, such as overtreatment [2], low situational awareness of planning staff, and limited knowledge of General Practitioners on mental care providers [3], impact this problem.

In this paper, we investigate its causes and possible solutions by analysing the mental healthcare system and creating a framework of this system showing the relationships between the factors that affect the problems of the long access times and low trust of patients. Using this framework, we develop a plan for effectively implementing these solutions. For this study, we build on the knowledge that is already used in other healthcare sectors and apply it to the mental healthcare sector. Additionally, we carefully consider human factors, which are especially important in a socially-oriented system like mental healthcare.

In Section II, we present related work from other healthcare subsystems. In Section III, we describe a summary of the causes of the long access times and the low trust of clients, as well as the relationships between these causes. In Section IV, we more closely investigate the six root causes and propose solutions to them. Finally, we conclude the work in Section V by proposing solutions and directions for further research.

II. Related work

In a study on a care program for older adults with dementia and depression in the United States, researchers showed that it is crucial for caregivers to build a relationship with their clients. This way, they can establish trust, which is important because it creates an openness that enables mental support. Additionally, a problem that appeared in this study was that staff used workarounds because not all systems were usable. Therefore, they sometimes needed to do duplicate work. Furthermore, scheduling was a complex task and did result in suboptimal schedules, which also result in a waste of time [4].

Scheduling challenges in combination with human factors are the subject of more studies. Guinery et al. [5] state that in service provision it is important to not solely aim for efficiency because this will increase 'failure demand'. Failure demand is unnecessary demand, so unnecessary work. Instead, they look at the whole system and suggest several access strategies, like pre-booked appointment scheduling or home visits. Different clinics can use different strategies because the choice for appointment strategies is always based on a trade-off. Guinery et al. also suggest that nurses can sometimes take over a task from a specialist to create more capacity. Further, in scheduling, it is important to take human factors into account. This way, human errors can be prevented. For instance, considering staff fatigue levels in scheduling can significantly reduce these levels [6]. Additionally, taking into account human factors, like the nurses' skills or their compatibility, can strengthen the teamwork, causing the staff to work more efficiently [7].

An algorithm for scheduling staff, in their case in operating rooms, while taking into consideration the staff's wellbeing, is designed by Roland et al. [8]. They show the possibility to simultaneously focus on financial and human resources. For scheduling, Discrete Event Simulation could be used as well. This technique can provide insights into the workload and quality of care, while still being a simulation instead of a test in a real environment. So, this is a cost-effective and safe way to evaluate changes in the planning strategy [9].

The aforementioned studies touch upon some important factors to consider when designing a health care system. We

 ISBN: 978-1-61208-941-6

should pay attention to them when investigating the causes of the complexities in the mental healthcare sector specifically. For instance, also in mental healthcare, staff having workarounds can indicate that the system is not designed properly. Additionally, solutions that are used in other healthcare sectors should also be investigated in the mental healthcare sector. For instance, the pooling of capacity that Guinery et al. [5] suggest regarding nurses, could also be a solution in mental healthcare. We use these lessons in the remainder of this paper to apply them to the mental healthcare sector.

A. Planning and scheduling in mental healthcare

Next to the literature on human factors in scheduling challenges, we also reviewed the literature on planning and scheduling techniques in the mental healthcare sector. A study by Carey et al. [10] in Australia suggests that letting clients determine when and how often they need an appointment can be equally effective and more efficient than the current system in which the practitioners determine the frequency and number of appointments. This patient-led appointment scheduling system seems to have the same health outcome, but with fewer appointments. Therefore, it can reduce access times. However, to know if these results also apply for the Netherlands, additional research is needed, because of the different culture and financing system. Additionally, this scheduling solution includes care-related considerations that should be taken into account.

Van den Berg-Vreeken [11] reduced access times for a mental health organisation in the Netherlands by creating a blueprint schedule to provide clarity regarding whether there are vacant appointment slots. Furthermore, the study modelled the process as a Markov Decision Process to determine which patient types should be assigned to which practitioner types. This is a useful method for organisations in which some practitioner types form a bottleneck, whereas the other types have sufficient capacity.

Other initiatives to reduce access times in mental healthcare include the study of Weaver et al. [12], in which the Toyota Production System was used to let employees cooperate in identifying improvement possibilities in the organisation. As a result, the number of phone calls needed before the first appointment was reduced, thereby reducing wasted capacity, which enabled shorter access times. Jones et al. [13] reduced access times by a so-called ‘triage day’. Instead of spending much time reviewing the referral letters to determine the treatment plan, the paper introduced the use of a triage day in which many real-life intakes are performed. Next to a reduction of the access times, this also reduced the no-show rate. The relation between these two variables was shown by Gallucci et al. [14], who observed that among clients with access times of at most seven weeks, reducing access times reduces the no-show probability. Orme and Boswell [15] in turn studied the factors that influence pre-intake drop-out rates, such as age, access time and gender.

B. Planning and scheduling in healthcare

In healthcare in general, more research is done on planning and scheduling. For instance, cyclic appointment schedules can be used to cope with a fluctuating demand of scheduled and unscheduled patients while balancing access and waiting times [16]. Further, distributing the capacity among clients with different urgencies is studied by Zhou et al. [17]. They used a nonlinear mixed-integer programming model to distribute scarce MRI capacity fairly while also considering revenue. This situation differs from the situation in mental healthcare organisations, because rejecting clients would harm their reputation. Patrick and Puterman [18] also consider patients with differing urgencies, but without the option of rejection. They show that in case the capacity is not much bigger than the average demand, scheduling patients at the latest acceptable moment, eventually in overtime, results in a schedule that manages the difference in urgency best with regard to the access times. However, if the capacity is much lower than the average demand, as is the case in many mental healthcare organisations, this will result in enormous overtimes. At an online operational level, Schütz and Kolisch [19] used an iterative dynamic programming model to decide whether to accept or reject a request while maximising the expected profit, which can also be the health outcome. The outcomes of the model were stored in a lookup table and used in a Discrete Event Simulation (DES) to evaluate the effects. Additionally, the authors compared the performance of the model to the performances of four heuristics.

Capacity allocation is also studied by Zhou et al. [20], who developed a multi-objective stochastic programming model for the allocation of ward capacity. They solved this model using DES and linearised the model to be solvable by several algorithms, among which is an ϵ-constraint algorithm. A method to manage a general practitioner’s workload was introduced by Zander et al. [21]. They used Integer Linear Programming models to determine which new patients a practitioner should accept to keep his expected workload balanced with his capacity, considering the age and number of past visits of the patients. Further, Elkhuizen et al. [22] studied reducing access times by first eliminating the backlog by adding temporary capacity and afterwards keeping the access times low by adding structural capacity.

Finally, Bikker et al. [23] reduced access times for cancer patients by designing blueprint schedules for doctors to allocate the capacity to the several treatment types. They did this by first optimising the schedule using an integer linear programming model that considered the static parameters and thereafter evaluating the schedule using DES to include the dynamic parameters.

C. Planning and scheduling in unstable systems

In the mental healthcare sector, demand structurally exceeds capacity, meaning the system is unstable. Other systems dealing with instability are the kidney allocation system and the US public housing system. Bandi et al. [24] created a method to determine the access times in these kinds of systems to give more clarity to patients and people that are waiting for a house. Bassamboo and Randhawa [25] designed an alternative to the First Come, First Served (FCFS) policy in overloaded systems. Based on the amount of time a customer already waited, they predict his willingness to wait longer. Using this information in a scheduling policy, they reduced the queue length, abandonment ratio and access time. Stolyar and Tezcan [26] also created a model to determine which customer to help next. Their model bases this choice on the customer-specific reward. To the best of our knowledge, no studies are performed on reducing access times in unstable systems in which balking of clients is not common.

To summarise, we found that in the mental healthcare sector only limited research is performed on planning and scheduling. In the healthcare sector in general, more research is done, but these studies assume stable systems or reject demand that exceeds the capacity. Therefore, they do not treat the difficulties that mental healthcare organisations are facing regarding the long waiting lists and access times. We have not found studies considering unstable systems when reducing access times.

III. The problem

To analyse the problems and their causes in mental healthcare, we created the problem cluster shown in Figure 1. The issues addressed in the figure are extracted from literature, the news and from our observations and conversations in a Dutch mental healthcare organisation. In the figure, we can see that all problems lead to the final problem of a low trust of clients, which is a serious issue, as Heiden et al. [4] show. When we go back in the causal chain, we see that this low trust is caused by the long access times and by the fact that a client is treated by many different therapists [27]. The causes of the long access times are overtreating [2] on the one hand and low situational awareness of planning staff on the other hand, because the low situational awareness causes a less efficient use of the available capacity. Additionally, the fact that a client meets many therapists is caused by the heterogeneity of the clients and the limited knowledge of General Practitioners on mental care providers [3]. Furthermore, both the long access times and the high number of therapists might be caused by procedures that are not easily usable [4], a cause which needs more research and can differ between mental healthcare organisations, and by "cherry-picking", which means that some mental care providers choose to only treat clients with easy problems while referring more complex cases to bigger organisations [28]. The different causal factors do also interrelate. For instance, we see that the fact that many therapists treat the same client causes work to be performed in duplicate, which increases the workload, which already is very high [29]. This in turn motivates staff to leave, which causes that clients again get a new therapist. In the end, we observe six root causes. These are the causes that we should take a closer look at.

IV. Solutions to the six root factors

To solve the challenges mental healthcare is facing, we should investigate the six root problems and generate solutions to them.

A. Overtreating

Overtreating occurs if clients are treated for an unnecessarily long time. It might be that there is no suitable treatment for the client, but therapists do not like giving up on their clients. In other cases, the therapist might continue treatment, while the client may already be recovered [2]. In this case, the therapist can be too careful and worried about a possible relapse of his client.

To reduce overtreating, firstly, it is important that therapists are aware of their possible tendency to overtreat. This can be done by comparing the average client cycle times of different therapists. In doing this, it is important to take into account differences between clients, because some disorders are easier to treat, while others are more complex and unpredictable. Additionally, in comparing therapists, it is important to consider privacy issues. Therapists that turn out to overtreat might feel attacked if it is noticed and this can affect their working ethos. So, it is important to create a safe environment in which therapists feel they can learn from each other and in which it is appreciated if they are open about their weaknesses. In an environment like that, development, improvement and growth can be realised. Secondly, reducing overtreatment can be realised by handing clients over to a General Practitioner when they are in a stable condition. To encourage this transfer, a close collaboration between the practitioner and the therapist is important [30].

B. Low situational awareness among planning staff

The low situational awareness among planning staff results in gaps in the practitioners' schedules. In multi-therapist practices, the therapists might need to notify the planning staff themselves when one of their clients has finished his treatment process. If the therapist does not immediately do this, no new client from the waiting list will be scheduled, which results in gaps, and therefore time waste. We can distinguish three levels within situational awareness: perception, comprehension and projection [31]. In this case, the main problem can be classified as a comprehension problem. The manager knows the relationship between the manual work, the planning and the gaps in the schedules, but just does not know the current situation. Therefore, the planning staff can also not project the current situation to the future, which is a projection problem. So, it is important to make sure the current situation is comprehended to enable planners to create a suitable schedule for the next weeks. This can be done, for instance, by creating a system that automatically notifies the planners if a therapist can take a new client. Another solution is to create reminders, for example in the form of a monthly email or a poster, to remind the practitioners to notify the planners. This can increase the situational perception of the practitioners. However, with this solution, there is a risk of alarm fatigue. Therefore, automation

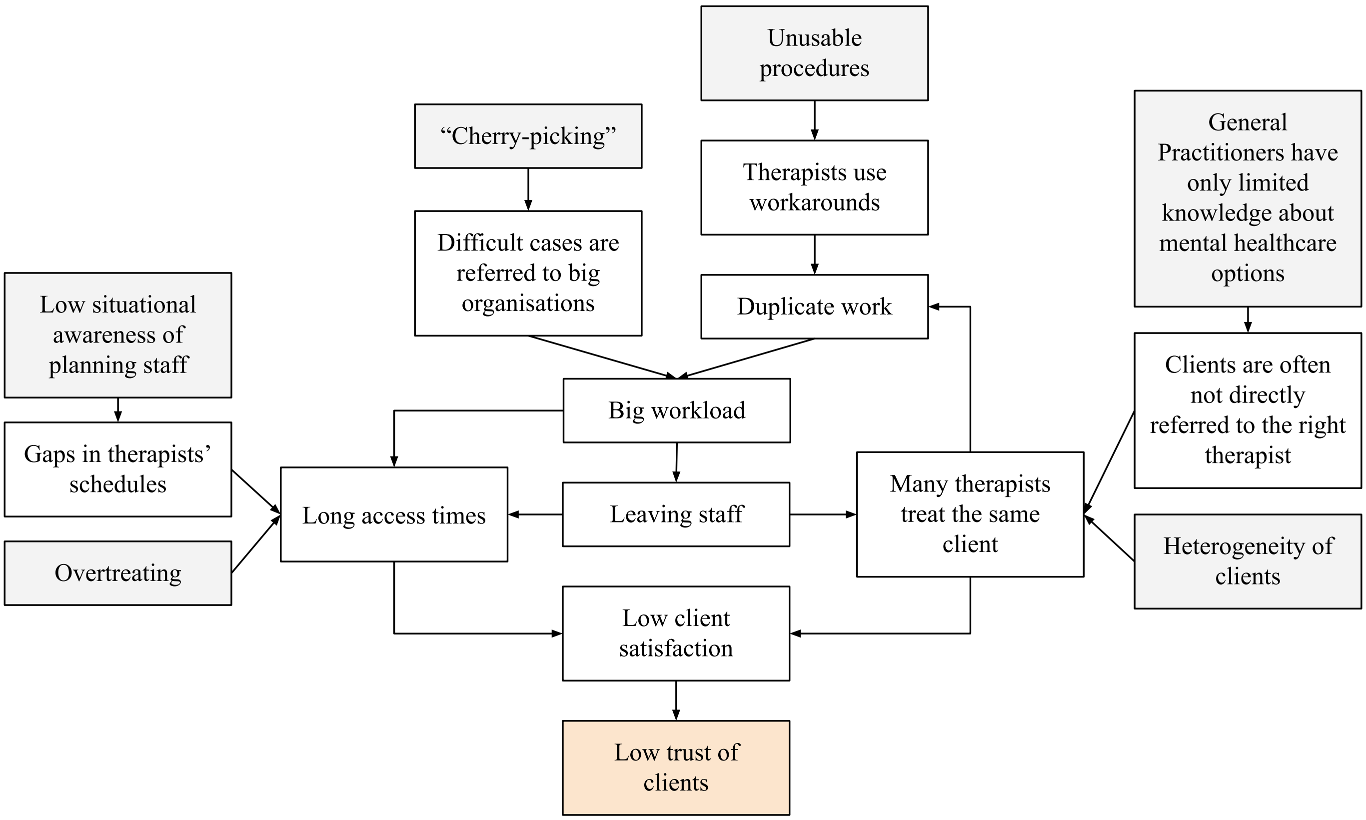


Figure 1. Mental healthcare problem cluster

is more likely to be effective. Nonetheless, the system should be developed with care and evaluation and participation of stakeholders should be part of the design process to ensure the system is effective.

C. "Cherry-picking"

The practice of cherry-picking among care providers can be explained by its financial attractiveness. Namely, up to the end of 2021, due to the national care funding system, short treatments with a high chance of success were more profitable than lengthy treatments with less or slower success. The advantage of this system was that care providers had an incentive to provide a high quality of care and limit overtreatment. However, it is hard to fairly benchmark care providers based on their success because mental diseases are all different and cover a wide range of complexities. Therefore, the chances of success differ much and it is not desirable that clients with more complex problems are not helped. To give providers an incentive to stop cherry-picking, the national care funding system should be adapted or legislation should require care providers to accept more clients with complex problems. The first, however, appeared to be difficult because objections related to cherry-picking were made against the national care funding system without success [32], because the Dutch Healthcare Authority stated that care providers are allowed to specialise in certain treatments. Nonetheless, since 2022 a new national care funding system is in place, which promises to make cherry-picking less attractive [33]. Forcing care providers to accept more clients with complex problems is difficult to implement because it would complicate the referrals that are needed when it turns out another provider offers a more suitable treatment for a client.

D. Unusable procedures

The fact that therapists might deliberately or instinctively use workarounds for procedures indicates that these procedures are not well-aligned with the needs of the therapists or clients. This misalignment can be observed, for instance, if procedures are performed on paper instead of on the computer or if procedures are even totally ignored. Some reasons for working on paper instead of in a computer program can be that the computer program is difficult to use or that it has flaws. Additionally, working on a computer can feel more impersonal, while working on paper seems to offer more flexibility [4]. Another workaround can be that therapists who use standardized questionnaires reword the questions [4]. This indicates that the questions might not be formulated correctly. This can either cause incorrect answers of clients if they misinterpret the question or it might hinder correct comparisons between clients of different therapists. These incorrect comparisons are then caused by the fact that researchers think the questions are standardized, whereas, in reality, the clients answered different questions because their therapists reformulated them. Additionally, workarounds can be used if the staff simply does not understand the procedures well enough to use them. All in all, it would be useful to further investigate which procedures are not usable enough and how considering human factors can improve the usability and usefulness of these procedures. This can either be done by adapting the procedures or by training staff. Although training staff can seem the easiest solution, it might just be extinguishing the fire. Creating an intuitive, useful system will be more effective if it is possible [34].

E. *Limited knowledge among General Practitioners*

If General Practitioners (GPs) do not know what mental healthcare providers have to offer, it is hard for them to refer their patients to the most suitable therapist. Therefore, clients might end up at the wrong therapist, where they will be on the waiting list for several weeks. After that time, they have an initial consult in which it turns out they should be referred to another therapist. There, the waiting process starts again. This is not desirable. We again find that a low situational awareness is at the heart of the problem. The perception of GPs is insufficient. To solve this issue, we could train GPs. However, the downside of this is that it improves the situation just temporarily, because the options for mental care are evolving and changing. Instead, in Limburg, a region in the Netherlands, a regional triage panel has been established. People from different backgrounds, ranging from several mental health organisations to the local government, together decide which organisation is most suitable to help the client [28]. This way, knowledge is combined, a client has a higher likeliness of being referred to the most suitable organisation and his access times at less suitable therapists are reduced.

F. *Heterogeneity of clients*

Mental difficulties are very heterogeneous. They cover a wide spectrum of symptoms and they evolve in many different ways. Therefore, it is not always clear which treatment will be successful. So, in the end, clients are treated by many different therapists. Each unsuccessful treatment lowers the clients' trust in the care system. In addition, the fact that they need to tell their situation repetitively to their new therapists is not beneficial for their satisfaction with the system. The reduction of trust should be minimised because trust is important in mental care [4]. If clients do not trust their therapists or their treatment, the treatment will be less successful, because trust is needed to encourage honesty and effort of the clients, which increase the chance of success.

However, the heterogeneity of clients is uncontrollable. Nonetheless, mental health specialists could develop so-called 'care paths'. These are standardised treatment plans for several different mental issues. Clients might not fit perfectly within these standardised paths, but the paths can still help in decision making. They can roughly show which therapy or medication a client should receive, after which the therapist can adapt the plan to the specific client. At a later stage, even artificial intelligence could be used to suggest care paths for a client. However, both with and without artificial intelligence, the risks of under- and overtrust by therapists should be acknowledged [35]. These risks could be dealt with, for instance, by motivating or forcing therapists to first come up with a plan themselves that they can later check against the standardised path. However, this increases the workload. Instead, therapists could be encouraged to first choose a care path, which saves time, and then adapt the care path to the needs of their specific client, which decreases the chance of overtrust. Whether this indeed pays off should be evaluated carefully.

V. CONCLUSION

The mental healthcare sector is coping with several complexities, like long access times and low trust of clients. In this paper, we outlined several causes and their interdependencies. The six root causes are overtreatment, low situational awareness of planning staff, cherry-picking, unusable procedures, limited knowledge among GPs, and heterogeneity of clients. These factors should be addressed to reduce workloads and access times and to increase clients' trust and satisfaction.

A. *System development*

To address the issues mentioned previously, it would be useful to create a system to improve situational awareness among therapists and planning staff. The system should keep track of the number of clients of therapists and it should notify the planning staff when they can assign a new patient to a therapist. Additionally, this system could compare the average client cycle times of different therapists to make therapists aware of their possible tendency to overtreat clients. Furthermore, standardized care paths, and eventually artificial intelligence, can be included in this system to support decision making. Hereby, the risk of over- and undertrust should be considered. Additionally, in designing such a system, it is important to let stakeholders participate in the design process and to prototype and evaluate the system repetitively to ensure that it is usable. This way the system is more effective and therapists are more likely to accept the system instead of developing workarounds that can cause duplicate work or errors.

B. *Collaboration*

Additionally, the mental health sector can benefit from collaboration between care providers. A regional triage panel can help to improve the success of the first referral of clients and a close collaboration between therapists and GPs can encourage a timely handover of the client from the therapist to the GP. In both approaches, trust among the stakeholders is crucial. Therefore, it is important they participate and think along in establishing these collaborations.

C. Future research

Some other aspects should be investigated further. It would, among others, be useful to investigate which procedures are not functioning properly. This can be done by observing whether staff uses workarounds. If procedures turn out to be unusable, projects can be started to improve the procedures. Hereby, user participation is important to ensure the procedures serve their purpose.

Finally, it would be useful to study whether and how the national care funding system could be further adapted to avoid cherry-picking, while still enabling to refer patients to other care providers if needed.

REFERENCES

[1] Vektis, ”Factsheet Wachttijdinformatie ggz [Factsheet Waiting time information ggz]”, https://www.vektis.nl/intelligence/publicaties/factsheet-wachttijdinformatie-ggz, retrieved: June, 2021.
[2] K. Kraaijeveld and R. Lomans, ”De ggz in 2025: Vergezicht op de geestelijke gezondheidszorg [”The ggz in 2025: Vision on the mental healthcare sector]”, 2020.
[3] F. Verhulst, ”Einde aan ’GGZ-jungle’: Nieuw platform moet zorg bij psychische klachten gaan stroomlijnen [End of the ’GGZ jungle’: New platform should streamline care for psychological complaints]”, BN DeStem, 2021.
[4] S. M. Heiden, R. J. Holden, C. A. Alder, K. Bodke, and M. Boustani, ”Human factors in mental healthcare: A work system analysis of a community-based program for older adults with depression and dementia”, Applied Ergonomics, vol. 64, pp. 27-40, 2017.
[5] J. Guinery, S. Brown, M. Berglund, and K. Scales, ”The influence of human factors on access and scheduling of primary care services. Advances in Production Management Systems”, Innovative and Knowledge-Based Production Management in a Global-Local World, pp. 547-554, 2014.
[6] R. -C. Lin, M. Y. Sir, E. Sisikoglu, K. Pasupathy, and L. M. Steege, ”Optimal nurse scheduling based on quantitative models of work-related fatigue”, IIE Transactions on Healthcare Systems Engineering, vol. 3(1), pp. 23-38, 2013.
[7] M. Hamid, R. Tavakkoli-Moghaddam, F. Golpaygani, and B. Vahedi-Nouri, ”A multi-objective model for a nurse scheduling problem by emphasizing human factors”, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, vol. 234(2), pp. 179-199, 2020.
[8] B. Roland, C. Di Martinelly, F. Riane, and Y. Pochet, ”Scheduling an operating theatre under human resource constraints”, Computers & Industrial Engineering, vol. 58(2), pp. 212-2020, 2010.
[9] S. M. Qureshi, N. Purdy, A. Mohani, and W. P. Neumann, ”Predicting the effect of nurse-patient ratio on nurse workload and care quality using discrete event simulation”, Journal of Nursing Management, vol. 27(5), pp. 971-980, 2019.
[10] T. Carey, S. Tai, and W. Stiles, ”Effective and efficient: Using patient-led appointment scheduling in routine mental health practice in remote Australia”, Professional Psychology: Research and Practise, vol. 44(6), pp. 405-414, 2013.
[11] S. van den Berg-Vreeken, ”A policy approximation of a Markov decision process for scheduling clients in an outpatient mental healthcare clinic”, 2021.
[12] A. Weaver et al., ”The Impact of System Level Factors on Treatment Timeliness: Utilizing the Toyota Production System to Implement Direct Intake Scheduling in a Semi-rural Community Mental Health Clinic”, The Journal of Behavioral Health Services & Research, vol. 40(3), pp. 294-305, 2013.
[13] E. Jones, C. Lucey, and L. Wadland, ”Triage: a waiting list initiative in a child mental health service”, Psychiatric Bulletin, vol. 24(2), pp. 57-59, 2000.
[14] G. Gallucci, W. Swartz, and F. Hackerman, ”Impact of the Wait for an Initial Appointment on the Rate of Kept Appointments at a Mental Health Center”, Psychiatric services, vol. 56(3), pp. 344-346, 2005.
[15] D. Orme and D. Boswell, ”The pre-intake drop-out at a community mental health center”, Community Mental Health Journal, vol. 27(5), pp. 375-379, 1991.
[16] N. Kortbeek et al., ”Designing cyclic appointment schedules for outpatient clinics with scheduled and unscheduled patient arrivals”, Performance Evaluation, vol. 80, pp. 5-26, 2014.
[17] L. Zhou, N. Geng, Z. Jiang, and X. Wang, ”Combining revenue and equity in capacity allocation of imaging facilities”, European Journal of Operational Research, vol. 256(2), pp. 619-628, 2017.
[18] J. Patrick and M. Puterman, ”Reducing Wait Times through Operations Research: Optimizing the Use of Surge Capacity”, Healthcare policy, vol. 3(3), pp. 75-88, 2008.
[19] H. Schütz and R. Kolisch, ”Capacity allocation for demand of different customer-product-combinations with cancellations, no-shows, and overbooking when there is a sequential delivery of service”, Annals of Operations Research, vol. 206(1), pp. 401-423, 2013.
[20] L. Zhou, N. Geng, Z. Jiang, and X. Wang, ”Multi-objective capacity allocation of hospital wards combining revenue and equity”, Omega (United Kingdom), vol. 81, pp. 220-233, 2018.
[21] A. Zander, S. Nickel, and P. Vanberkel, ”Managing the intake of new patients into a physician panel over time”, European Journal of Operational Research, vol. 294(1), pp. 391-403, 2021.
[22] S. Elkhuizen, S. Das, P. Bakker, and J. Hontelez, ”Using computer simulation to reduce access time for outpatient departments”, Quality and Safety in Health Care, vol. 16(5), pp. 382-386, 2007.
[23] I. Bikker, N. Kortbeek, R. van Os, and R. Boucherie, ”Reducing access times for radiation treatment by aligning the doctor’s schemes”, Operations Research for Health Care, vol. 7, pp. 111-121, 2015.
[24] C. Bandi, N. Trichakis, and P. Vayanos, ”Robust Multiclass Queuing Theory for Wait Time Estimation in Resource Allocation Systems”, Management Science, vol. 65(1), pp. 152-187, 2019.
[25] A. Bassamboo and R. Randhawa, ”Scheduling Homogeneous Impatient Customers”, Management Science, vol. 62(7), pp. 2129-2147, 2015.
[26] A. Stolyar and T. Tezcan, ”Shadow-routing based control of flexible multi-server pools in overload”, Operations Research, vol. 59(6), pp. 1427-1444, 2011.
[27] J. Maas, ”Vertrouwen is hét instrument van de hulpverlener [Trust is the main instrument of the health care provider]”, https://www.zorgwelzijn.nl/vertrouwen-is-het-instrument-van-de-hulpverlener/, retrieved: March, 2022.
[28] Weg van de Wachtlijst, ”Regionale triagetafel voor verwijzingen naar de specialistische GGZ [Regional triage table for referrals to specialised GGZ]”, https://wegvandewachtlijst.nl/regionale-triagetafel-voor-verwijzingen-naar-de-specialistische-ggz, retrieved: June, 2021.
[29] M. van Wijk, ”Arbeidsmarktprofiel van zorg en welzijn [Job market profile of the healthcare sector]”, https://www.cbs.nl/nl-nl/longread/statistische-trends/2020/arbeidsmarktprofiel-van-zorg-en-welzijn?onepage=true, retrieved: March, 2022.
[30] R. Meliesie, ”‘Warme overdracht’: continuïteit van zorg, afgestemd op de patient [’Warm transfer’: continuity of care, tailored to the patient]”, https://wegvandewachtlijst.nl/warme-overdracht-continuiteit-van-zorg-afgestemd-op-de-patient, retrieved: June, 2021.
[31] M. R. Endsley, ”Situational awareness analysis and measurement”, Lawrence Erlbaum Associates., 2000.
[32] J. M. T. Hautvast, ”Beslissing op bezwaar - ongegrond (1144449) [Decision on appeal - unfounded (1144449)]”, 2019
[33] H. Maassen, ”Bekostiging ggz gaat op de schop [GGz funding is going to be changed]”, https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/bekostiging-ggz-gaat-op-de-schop.htm, retrieved: March, 2022.
[34] A. L. Russ et al., ”The science of human factors: separating fact from fiction”, BMJ Quality & Safety, vol. 22(10), pp. 802-808, 2013.
[35] O. Asan, A. E. Bayrak, and A. Choudhury, ”Artificial intelligence and human trust in healthcare: focus on clinicians”, Journal of Medical Internet Research, vol. 22(6), 2020.